# Abstract

This thesis is presented in two parts. In the first part, we study a family of models of random partial orders, called classical sequential growth models, introduced by Rideout and Sorkin as possible models of discrete space-time. We analyse a particular model, called a random binary growth model, and show that the random partial order produced by this model almost surely has infinite dimension. We also give estimates on the size of the largest vertex incomparable to a particular element of the partial order. We show that there is some positive probability that the random partial order does not contain a particular subposet. This contrasts with other existing models of partial orders. We also study "continuum limits" of sequences of classical sequential growth models. We prove results on the structure of these limits when they exist, highlighting a deficiency of these models as models of space-time.

In the second part of the thesis, we prove some correlation inequalities for mappings of rooted trees into complete trees. For $T$ a rooted tree we can define the proportion of the total number of embeddings of $T$ into a complete binary tree that map the root of $T$ to the root of the complete binary tree. A theorem of Kubicki, Lehel and Morayne states that, for two binary trees with one a subposet of the other, this proportion is larger for the larger tree. They conjecture that the same is true for two arbitrary trees with one a subposet of the other. We disprove this conjecture by analysing the asymptotics of this proportion for large complete binary trees. We show that the theorem of Kubicki, Lehel and Morayne can be thought of as a correlation inequality which enables us to generalise their result in other directions.

# Contents

# List of Tables

# List of Figures

To my parents, for all their support.

# Acknowledgements

Firstly, I would like to thank Professor Graham Brightwell for his unfaltering supervision. He introduced me to the problems studied here and he has been a constant inspiration throughout my studies. I also extend my thanks to the whole of the Mathematics Department at LSE. I have really appreciated the informal atmosphere around the department, and the friendly nature of all its members. I am sure that this is owing, in part, to the informal workshops and discussions at the round table—long may they continue. Particular thanks go to Dave Scott and Jackie Everid for their administrative support, Mark Baltovic for his IT expertise, and my fellow research students for sharing the highlights and frustrations of research and for our many conversations on a multitude of topics, a most welcome distraction.

Away from the university, my thanks go to all my friends living in and around London who have made my time here such an enjoyable one. Special thanks go to Owen Jones, Katrine Reimers and Esme Tarasewicz for being wonderful housemates, and to Joe Walton and Nick Sample for their generosity and friendship.

My thanks, as always, to my parents and my brother, Matthew, for their loving support and encouragement.

Finally, I thank the Engineering and Physical Sciences Research Council, the London School of Economics and Political Science, and the Department of Mathematics for their financial support over the last four years.
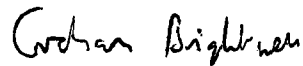
London, September 2005

# Statement of originality

I declare that the work described in this thesis is wholly my own, except for the work in Chapter 3. The work described in Chapter 3 was carried out in conjunction with my supervisor, Professor Graham Brightwell and was worked on in equal proportion by myself and Professor Brightwell.

Nicholas Georgiou (candidate)

Graham Brightwell (supervisor)

# Summary

This thesis covers two areas in probabilistic combinatorics, specifically the combinatorics of partially ordered sets. Problems and areas of study in probabilistic combinatorics broadly fall into one of two classes. The first class contains problems of a deterministic nature, which are particularly suited to some application of probabilistic methods or techniques. The second class contains problems that are themselves of a probabilistic nature. We cover problems from both classes.

In the first part of the thesis we investigate a family of random models of partial orders, called classical sequential growth models. We study in detail the simplest non-trivial model from the family and analyse the partial orders it produces. We also study "continuum limits" of sequences of classical sequential growth models, proving that particular sequences of these models do have continuum limits. We also prove some results about the continuum limit of a general sequence of classical sequential growth models, when it exists.

In the second part of the thesis we look at enumeration of embeddings of trees into complete trees, which can be motivated by a partial-order variant of the best secretary problem. We show that a monotone property of binary trees that was conjectured to hold for arbitrary trees does not hold in general. We show that the monotonicity on binary trees is an example of a correlation inequality on a certain lattice, and using this we can prove generalisations in other directions.

# Part I

# Classical sequential growth models

In this part we study a family of models of random partial orders, called classical sequential growth models, introduced by Rideout and Sorkin [24]. These models were proposed as possible models for discrete space-time, since they are the only models satisfying certain desirable physical-looking conditions. In particular, we will analyse the simplest non-trivial model from the family, and we will also define and study a particular limit of a sequence of classical sequential growth models.

In Chapter 1 we give a full description of the family of models and a brief summary of the results in [24], explaining the physical-looking conditions imposed by Rideout and Sorkin, and noting that a particular model from the family can be specified by a sequence of non-negative constants.

In Chapter 2 we study in detail the particular model called a random binary growth model, showing that a random poset produced by the model almost surely has infinite (poset) dimension. This shows that, despite the simple description of the model, the random poset it produces has a complex structure. We give estimates for bounds on the size of an up-set of a particular element and show that every element in the random infinite poset is incomparable to only finitely many others. We also present a specific poset that, with some positive probability, is not contained in the random poset produced by the model. This contrasts the model with other random models of partial orders, for example the random graph order, which contains any specific poset almost surely.

In Chapter 3 we study the continuum limits of sequences of classical sequential growth models. Rideout and Sorkin [25] have provided computational evidence suggesting that particular sequences of models have a continuum limit. We formalise their results by defining what a continuum limit is, and we show that if a sequence has a continuum limit then it must be an almost-semiorder. Using some results of Pittel and Tungol [23] on random graph orders, we prove that the continuum limit of a sequence of random graph orders, when it exists, is a random semiorder. We also present some new results on classical sequential growth models. This chapter

describes work carried out in conjunction with my supervisor, Professor Graham Brightwell.

# Chapter 1

# Introduction

We study a family of models of random partial orders, called *classical sequential growth models*, introduced by Rideout and Sorkin [24]. Each model is defined on the (labelled) vertex set $\mathbb{N}$, which we will always take to include 0. Any model can be restricted to $[n] = \{0, 1, 2, \ldots, n\}$ and regarded as a model of random finite posets. The model starts with a poset of one element (labelled 0), and grows in stages. At stage $n = 1, 2, \ldots$, vertex $n$ is added to the existing poset, $P_{n-1}$, by placing $n$ above some choice of vertices of $P_{n-1}$. The poset $P_n$ is defined on vertex set $[n]$ by taking the transitive closure of the existing and added relations. This is called a *transition* from $P_{n-1}$ to $P_n$, written $P_{n-1} \to P_n$. The models are random, so each transition occurs with some probability. These *transition probabilities* are fixed and depend on the particular model. Let $\mathbb{P}(P_{n-1} \to P_n)$ denote the probability of transition $P_{n-1} \to P_n$ occurring.

Rideout and Sorkin then impose four conditions on the transition probabilities, with the aim of giving the model physical meaning. They call these conditions: internal temporality, discrete general covariance, Bell causality and Markov sum. The first and last conditions are implicit in the mathematical approach to random partial orders, namely that the labelling of a poset is natural (can be extended to the $<$ order on the natural numbers), and that the model is indeed "random"

(at each stage $n$ and for any fixed $P_{n-1}$ the sum of probabilities over all possible transitions $P_{n-1} \to P_n$ must be equal to 1). *Discrete general covariance* states that the probability of producing a particular poset should not depend on the labelling of the poset, that is, given two different sequences of transitions, $(P_i \to P_{i+1})$ and $(Q_i \to Q_{i+1})$ which produce the isomorphic posets $P_n$ and $Q_n$, the products

$$\prod_{i=0}^{n-1} \mathbb{P}(P_i \to P_{i+1}) \quad \text{and} \quad \prod_{i=0}^{n-1} \mathbb{P}(Q_i \to Q_{i+1})$$

must be equal. So, for example, discrete general covariance immediately implies that any two transitions from $P_{n-1}$ to isomorphic posets $P_n$ and $P'_n$ have the same transition probability $\mathbb{P}(P_{n-1} \to P_n) = \mathbb{P}(P_{n-1} \to P'_n)$. *Bell causality* is a condition on ratios of transition probabilities. (Note that in [24] Rideout and Sorkin only study "generic" models, meaning that all transition probabilities are non-zero, in order to make sense of this condition.) Given a particular poset $P$, and any two transitions $P \to P'$, $P \to P''$ which add the new element $n$, let $S$ be the set of all elements which are incomparable with $n$ in both $P'$ and $P''$. Let $Q$ be the poset formed from $P$ by removing all the elements of $S$ (and obsolete relations), and define $Q'$ and $Q''$ similarly. Then, Bell causality states that

$$\frac{\mathbb{P}(P \to P')}{\mathbb{P}(P \to P'')} = \frac{\mathbb{P}(Q \to Q')}{\mathbb{P}(Q \to Q'')},$$

the idea being that, since the new element is not placed above any of the elements of $S$ in either transition, the presence of the set $S$ should not affect the ratio of the transition probabilities.

A particular model is specified by a sequence $\mathbf{t} = (t_0, t_1, \ldots)$ of non-negative constants. The random poset is defined as the transitive closure of a directed random graph $G_{\mathbf{t}}$ on $\mathbb{N}$ in which all arcs go from a lower numbered vertex to a higher. The arcs are selected sequentially, considering each vertex $n$ in turn and choosing the set $D_n \subseteq [n-1]$ of vertices sending an arc to $n$; the probability that $D_n$ is equal to a set $D$ being proportional to $t_{|D|}$, so that

$$\mathbb{P}(D_n = D) = \frac{t_{|D|}}{\sum_{j=0}^{n} \binom{n}{j} t_j}.$$

A model defined according to this description is called a *classical sequential growth model*. Rideout and Sorkin show that these models are the only generic models satisfying their conditions. It is an easy exercise to check that these models do indeed satisfy the four conditions; for example, Bell causality holds essentially because the relative probability that element $n$ selects a set $D$, defined as $t_{|D|}$, is independent of $n$.

Varadarajan and Rideout [31] and Dowker and Surya [12] have studied the situation where the transition probabilities are allowed to be zero. The Bell causality condition becomes a condition on products of transition probabilities and the type of models that satisfy the conditions are very similar to the generic models described here.

The family of classical sequential growth models also contains models of *random graph orders*. A random graph order $P_{n,p}$ is defined as follows. The ground set of $P_{n,p}$ is the set $\{0, 1, \ldots, n-1\}$. For each pair of vertices $i < j$ the relation $(i,j)$ is introduced with probability $p$. The poset $P_{n,p}$ is then the transitive closure of these relations. Random graph orders were introduced by Albert and Frieze [1] and have been studied further by Bollobás and Brightwell [7, 8, 9] and Simon, Crippa and Collenberg [27]. The area is covered in the survey of random partial orders by Brightwell [10]. A classical sequential growth model defined by sequence **t** where $t_i = t^i$ for all $i$, and $t = p/(1-p)$, will after stage $n-1$ produce a random graph order $P_{n,p}$.

In the following chapter, we concentrate on the model where the sequence **t** is $(0, 0, 1, 0, \ldots)$, i.e., where all $t_i$ are zero except $t_2$. This means that $|D_n| = 2$ for each vertex $n$. We say that $n$ *selects* the two vertices in $D_n$. So, in this model each vertex $n$ selects two vertices chosen uniformly at random from the set $[n-1]$. We assume that we start with the vertices 0 and 1 incomparable with probability 1 and then add vertices $n = 2, 3, \ldots$ according to the model. (So, for example, $D_2 = \{0, 1\}$ with probability 1.) We call this model a *random binary growth model* and call the

random poset it produces a *random binary order*.

This is the simplest interesting model; the model defined by **t** with $t_0$ non-zero and $t_i$ equal to zero for $i \geq 1$ produces an infinite antichain ($D_n = \emptyset$ with probability 1, for all $n$), and the model defined by **t** with $t_0$ and $t_1$ non-zero and $t_i$ equal to zero for $i \geq 2$ produces a forest of infinitely many infinite trees, where each vertex is an upper cover of exactly one other vertex and a lower cover of infinitely many other vertices. These are called the "dust universe" and "forest universe", respectively, in [24].

The random binary growth model also has potential applications in computer science. Under the name of a *random binary recursive circuit*, the random binary order has been studied by Mahmoud and Tsukiji [20, 21], Tsukiji and Xhafa [30] and Arya, Golin and Mehlhorn [4]. These papers typically focus on the "depth" of the circuit or the number of "outputs" of the circuit. These are considered as important parameters in a computer science setting; however, they correspond to the height of the random binary order and the number of maximal elements of the random binary order, which are not particularly interesting parameters of a partial order. Here we will consider parameters that are more interesting from a combinatorial viewpoint, but these will probably not have useful analogues in the recursive circuit formulation.

The random binary growth model is essentially the same as any other model with $t_3 = t_4 = \ldots = 0$ since for large $n$ the number of 2-element subsets of $[n-1]$ is significantly greater than the number of 1-element subsets and so the probability of $n$ selecting just one vertex (or no vertices) is very small in comparison to the probability of $n$ selecting two vertices. Therefore the results in the following chapter will carry over easily to such models.

# Chapter 2

# The random binary growth model

Recall that the random binary growth model is defined as follows. Start with elements 0 and 1 incomparable; then each element $n = 2, \ldots$ selects two elements uniformly at random from $[n-1]$, and we take the transitive closure. We will denote the *random binary growth model* by $\mathcal{B}_2$ and the *random binary order* it produces by $B_2$. We write $B_2[n]$ for the restriction of $B_2$ to $[n]$ and $B_2[n_1, n_2]$ for the restriction of $B_2$ to $[n_1, n_2] = \{x \in \mathbb{N} : n_1 \le x \le n_2\}$.

The random binary order $B_2$ is a sparse order; each vertex $n$ has at most 2 lower covers since $x$ is a lower cover of $n$ if and only if it is selected by $n$ and is not below the other vertex $y$ selected by $n$. This means the Hasse diagram of $B_2[n]$ has at most $2n$ edges. Also, as we now show, the expected width (i.e., the expected size of the largest antichain) of $B_2[n]$ increases with $n$. A vertex $x$ in $B_2[n]$ is maximal if and only if all vertices $y = x+1, x+2, \ldots, n$ do not select $x$, so

$$\mathbb{P}(x \text{ is maximal in } B_2[n]) = \prod_{y=x+1}^{n} \left(1 - \frac{2}{y}\right) = \prod_{y=x+1}^{n} \frac{y-2}{y} = \frac{x(x-1)}{n(n-1)}$$

and so the expected number of maximal elements is

$$\frac{1}{n(n-1)} \sum_{x=2}^{n} x(x-1) = \frac{1}{n(n-1)} \left(\sum_{x=1}^{n} x^2 - \sum_{x=1}^{n} x\right)$$

$$= \frac{1}{n(n-1)} \left(\frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2}\right) = (n+1)/3.$$

(In fact, this is shown in [20], where Mahmoud and Tsukiji also show that the number of maximal elements of $B_2[n]$ tends in distribution to a normal random variable with mean $n/3$ and variance $4n/45$.) The maximal elements form an antichain, so the expected width of $B_2[n]$ is at least $(n+1)/3$.

However, the number of minimal elements is always 2, since only 0 and 1 are minimal. Moreover, the expected number of minimal elements of $B_2[n_1, n_2]$, for $n_1 \geq 2$, is bounded above by $n_1$ as $n_2$ tends to infinity. Indeed, a vertex $x$ in $B_2[n_1, n_2]$ is minimal if and only if it selects both vertices from $[n_1 - 1]$, and the probability of this is $\binom{n_1}{2}/\binom{x}{2} = n_1(n_1 - 1)/x(x - 1)$. Summing over $x$ from $n_1$ to $n_2$ gives the expected number of minimal elements equal to $n_1 - n_1(n_1 - 1)/n_2$.

In Section 2.1 we study the dimension of $B_2$. The dimension of a poset $P$ on ground set $X$ is the minimum number of linear orders on the set $X$ whose intersection is equal to $P$. In other words, the minimum number of linear orders $L_i$ such that $x < y$ in $P$ if and only if $x < y$ in $L_i$ for all $i$. An equivalent definition is that the dimension of $P$ is the smallest $d$ such that $P$ can be embedded into $\mathbb{R}^d$, where $\mathbb{R}^d$ is the $d$-dimensional Euclidean space with ordering $(x_1, \ldots, x_d) < (y_1, \ldots, y_d)$ in $\mathbb{R}^d$ if $x_i < y_i$ in $\mathbb{R}$, for all $i = 1, \ldots, d$. (The equivalence can be easily proven; the essential observation is that the linear orders on $X$ correspond to the coordinate-wise orderings of the embedded points in $\mathbb{R}^d$.) Since $B_2$ is sparse, one might suppose there to be a relatively simple structure to $B_2$. However, we show this is not the case in so much as showing that $B_2$ has infinite dimension, almost surely. Using standard notation (see, e.g., [29]), we write $P(1, 2; m)$ for the subposet of the subset lattice formed by the 1-element and 2-element subsets of the $m$-element set $\{1, \ldots, m\}$ ordered by inclusion. Spencer [28] proved that the dimension of $P(1, 2; m)$ is greater than $\log_2 \log_2 m$, so we show that $B_2$ has infinite dimension, almost surely, by showing it contains a copy of $P(1, 2; m)$ as a subposet, for each $m$, almost surely. This is done by counting (and bounding the expected number of) certain "paths" in $B_2$ (the "paths" in $B_2$ are exactly the paths in the

directed random graph $G_t$).

In Section 2.2 we study the sizes of up-sets in $B_2[n]$ and, related to this, the number of elements in $B_2$ incomparable with an arbitrary element. Although $B_2$ is sparse, we show that for all $r$ the number of elements incomparable with $r$ is finite. In particular, this implies that $B_2$ does not contain an infinite antichain, almost surely. Moreover, for any classical sequential growth model defined by sequence $\mathbf{t}$ where $t_i \neq 0$ for some $i \geq 2$, the same result is true, that the random poset produced does not contain an infinite antichain, almost surely.

We use the differential equation method of Wormald [32, 33] which specifies when and how a discrete Markov process can be closely approximated by the solution to a related differential equation. We prove a version of Wormald's theorem which makes explicit the errors in the approximation. We use this result to analyse the growth of the up-set of an arbitrary point. For a fixed point $r$, write $U_r^{[n]}$ for the set of elements above $r$ in the finite poset $B_2[n]$. We can think of "growing the poset" by increasing $n$. Then $|U_r^{[n]}|$, which depends on $n$, can be considered as a Markov process. Using this "differential equation method", we give good estimates on $|U_r^{[n]}|$ for particular values of $n$, and show that there exists an $n = n(r)$ such that $I_r \subseteq [n]$. Here, $I_r$ is the set of vertices greater than $r$ which are incomparable with $r$. So, for fixed $r$, there are no vertices greater than $n$ incomparable to $r$, and so the number of vertices incomparable with $r$ is finite. We provide two similar proofs, one giving bounds for a typical $r$, and one giving bounds for all but finitely many $r$.

Is the fact that $P(1, 2, m)$ is almost surely contained in $B_2$ a special case of something more general? Is it possible, as in the case of random graph orders, that *every* finite poset is contained in $B_2$, almost surely? In Section 2.3 we show that this is not the case. We use our result from Section 2.2, that there is an $n = n(r)$ such that for all but finitely many $r$, there are no vertices greater than $n$ incomparable with $r$. So, we know that if two elements in $B_2$ have labels with a large enough difference then they must be comparable. We construct a poset which if contained

in $B_2$ must have two elements whose labels have large difference. Combining these two results, we provide an example of a poset not contained in $B_2$ (or rather, there is a positive probability that $B_2$ does not contain the poset).

## 2.1   The dimension of $B_2$

We write $P(1, 2; m)$ for the subposet of the subset lattice formed by the 1-element and 2-element subsets of the set $\{1, \ldots, m\}$ ordered by inclusion. For a particular vertex $r$, let $U_r$ be the set of all vertices above $r$ in $B_2$ and let $U_r^{[t]}$ be the set of all vertices above $r$ in $B_2[t]$. Denote by $T_k$ the hitting time of the event $|U_r| = k$, i.e., the smallest $t$ such that $|U_r^{[t]}| = k$, and the waiting time between events $|U_r| = k - 1$ and $|U_r| = k$ by $W_k$, so that $T_{k+1} = T_k + W_{k+1}$. We include the point $r$ in $U_r$ so that $T_1 = r$.

We now show that, for every $m$, there exists a copy of $P(1, 2; m)$ in $B_2$, almost surely. This is enough to show that $B_2$ almost surely has infinite dimension, since $\dim P(1, 2; m) \geq \log_2 \log_2 m$ (see [28]).

In fact, we will prove a stronger result, that for each $m$ there exists an $r_0$ such that the probability of there being a copy of $P(1, 2; m)$ in $B_2[r, 2r^{7/5}]$ is greater than $3/5$ for all $r \geq r_0$.

We use the following lemma to find a copy of $P(1, 2; m)$ in $B_2[r, 2r^{7/5}]$.

**Lemma 2.1.** *For any $m$, and any $n_1 < n_2$, if we have sets $X = \{x_1, \ldots, x_m\} \subseteq [n_1, n_2]$, $Y = \{y_1, \ldots, y_M\} \subseteq [n_1, n_2]$, where $M = \binom{m}{2}$, and the following conditions hold*

   *(i) the points in $X$ are incomparable in $B_2[n_1, n_2]$,*

   *(ii) for each pair of points $x_i, x_j$ in $X$ there is exactly one $y_k$ in $Y$ which is above these points and no others in $X$ (according to the order $B_2[n_1, n_2]$),*

*then $X \cup Y$ is a copy of $P(1,2;m)$ in $B_2[n_1, n_2]$ where $X$ is the set of minimal elements and $Y$ is the set of maximal elements.*

**Proof.** Let $<$ be the order on $B_2[n_1, n_2]$. To show that $X \cup Y$ is a copy of $P(1,2;m)$ we need to show that the only relations are those described by condition (ii). That is, that there are no relations of the form $x_i < x_j$, $y_k < y_l$ or $y_k < x_i$.

By condition (i) there are no relations of the form $x_i < x_j$. So, suppose there exists some relation $y_k < y_l$. Since $|Y| = M = \binom{m}{2}$, condition (ii) implies that there exists a pair $x_i, x_j$ with $x_i, x_j < y_k$. But then $x_i, x_j < y_l$ contradicting condition (ii). Suppose there exists some relation $y_k < x_i$. Then by condition (ii) there exists some $y_l$ with $x_i < y_l$. But then $y_k < y_l$ which leads to a contradiction as above.

So, $X \cup Y$ is a copy of $P(1,2;m)$ and $X$ is the set of minimal elements and $Y$ is the set of maximal elements. $\square$

**Proposition 2.2.** *For every $m$, there exists an $r_0$ such that the probability of there being a copy of $P(1,2;m)$ in $B_2[r, 2r^{7/5}]$ is greater than $3/5$ for all $r \geq r_0$.*

**Proof.** We will prove the result as follows. Assume that $m$ is fixed, $r_0$ is sufficiently large and $r \geq r_0$. First we find a set of points that satisfies condition (i) of Lemma 2.1 with some high constant probability. Because of the sparsity of $B_2$, it is easy to find this set. Here we will take the points $r, r+1, \ldots, r+m-1$. These points will form the minimal elements of a copy of $P(1,2;m)$. We then grow the poset up to size $r^{7/5}$, keeping track of the sizes of the up-sets of these chosen minimal points. The value $r^{7/5}$ is chosen so that the up-sets are large enough, but their pairwise intersection is still an insignificant fraction of the whole up-set. This means that the set of points above one and only one of the minimal points is reasonably large. The bulk of the proof is in showing this. Finally, we grow the poset up to size $2r^{7/5}$ to find the points satisfying condition (ii) of Lemma 2.1. Indeed, we look for points in $[r^{7/5}+1, 2r^{7/5}]$ selecting a pair of points from each pair of "exclusive up-sets". Because the sizes of the exclusive up-sets are known, we can show that the probability of finding these

points is at least some constant probability. We then apply Lemma 2.1 to obtain the result.

Following this scheme, where $m$ is fixed, $r_0$ is sufficiently large and $r \geq r_0$, consider the points $r, r + 1, \ldots, r + m - 1$. We attempt to find a copy of $P(1, 2; m)$ in which these are the minimal elements. We have,

$$\mathbb{P}(r, r+1, \ldots, r+m-1 \text{ are incomparable}) = \prod_{i=1}^{m-1} \frac{\binom{r}{2}}{\binom{r+i}{2}} \geq 9/10 \quad \text{for } r_0 \geq 20m^2.$$

Now grow the poset by adding points up to $n = r^{7/5}$. We consider the growth of the set $U_r$. We calculate the expected waiting time $\mathbb{E}W_{k+1}$ as follows. Suppose $T_k = t$, then since $W_{k+1}$ always takes integer values greater than or equal to 1 we have

$$\mathbb{E}W_{k+1} = 1 + \sum_{j=1}^{\infty} \mathbb{P}(W_{k+1} > j) = 1 + \sum_{j=1}^{\infty} \prod_{l=1}^{j} \frac{\binom{t+l-k}{2}}{\binom{t+l}{2}}$$

and using the inequalities $1 - x \leq e^{-x}$ and $\int_a^{b+1} f(x)dx \leq \sum_{j=a}^{b} f(j) \leq \int_{a-1}^{b} f(x)dx$, for $f$ decreasing, we have

$$\mathbb{E}W_{k+1} = 1 + \sum_{j=1}^{\infty} \prod_{l=1}^{j} \frac{\binom{t+l-k}{2}}{\binom{t+l}{2}} \leq 1 + \sum_{j=1}^{\infty} \left( \prod_{l=1}^{j} \left( \frac{t+l-k}{t+l} \right) \right)^2$$

$$\leq 1 + \sum_{j=1}^{\infty} \exp\left( -2 \sum_{l=1}^{j} \frac{k}{t+l} \right)$$

$$\leq 1 + \sum_{j=1}^{\infty} \exp\left( -2k \int_1^{j+1} \frac{1}{t+l} dl \right)$$

$$= 1 + \sum_{j=1}^{\infty} \left( \frac{t+1}{t+j+1} \right)^{2k}$$

$$\leq 1 + (t+1)^{2k} \int_0^{\infty} \frac{1}{(t+j+1)^{2k}} dj$$

$$= 1 + \frac{(t+1)^{2k}}{(t+1)^{2k-1}} \frac{1}{2k-1}.$$

That is,

$$\mathbb{E}(W_{k+1}|T_k) \leq 1 + \frac{T_k+1}{2k-1}.$$

So, we have

$$\mathbb{E}T_{k+1} = \mathbb{E}T_k + \mathbb{E}W_{k+1} \leq \mathbb{E}T_k + \left( 1 + \frac{\mathbb{E}T_k+1}{2k-1} \right) = \frac{2k}{2k-1}(\mathbb{E}T_k + 1), \qquad (2.1)$$

which by induction on $k$ gives

$$\mathbb{E}T_{k+1} \leq \left(2^{2k}/\binom{2k}{k}\right)r + 2k. \tag{2.2}$$

Using Stirling's approximation we have

$$\binom{2k}{k} \geq \frac{\sqrt{2\pi}(2k)^{2k+1/2}e^{-2k+1/(24k+1)}}{(\sqrt{2\pi}k^{k+1/2}e^{-k+1/12k})^2} \geq \frac{2^{2k+1/2}e^{1/(24k+1)}}{\sqrt{2\pi}k^{1/2}e^{1/6k}}, \quad \text{for } k \geq 1,$$

so $\mathbb{E}T_{k+1} \leq \sqrt{\pi}e^{1/6k-1/(24k+1)}\sqrt{k}r + 2k$, for $k \geq 1$. For $k \geq 2$, $\sqrt{\pi}e^{1/6k-1/(24k+1)} \leq 2$

and using (2.2) we have $\mathbb{E}T_2 \leq 2r + 2$, so $\mathbb{E}T_{k+1} \leq 2r\sqrt{k} + 2k$ and so

$$\mathbb{E}T_k \leq 2r\sqrt{k} + 2k. \tag{2.3}$$

If we similarly define $U_{r+i}, T_k^{(i)}, W_k^{(i)}$ for $r+i$, $i = 1, \ldots, m-1$ and write $T_k^{(0)}$ for $T_k$, then we have $T_1^{(i)} = r + i$, giving equations

$$\mathbb{E}T_{k+1}^{(i)} \leq \frac{2k}{2k-1}(\mathbb{E}T_k^{(i)} + 1), \tag{2.4}$$

$$\mathbb{E}T_{k+1}^{(i)} \leq \left(2^{2k}/\binom{2k}{k}\right)(r+i) + 2k, \tag{2.5}$$

$$\mathbb{E}T_k^{(i)} \leq 2(r+i)\sqrt{k} + 2k, \tag{2.6}$$

corresponding to equations (2.1),(2.2) and (2.3).

For $r_0 \geq m$ we have $r + i \leq r + m \leq 2r$, so (2.6) becomes

$$\mathbb{E}T_k^{(i)} \leq 4r\sqrt{k} + 2k, \quad i = 0, \ldots, m-1.$$

So, recalling that $n = r^{7/5}$, we have

$$\mathbb{P}(|U_r^{[n]}| < r^{3/4}) = \mathbb{P}(T_{r^{3/4}} > n) \leq \mathbb{E}T_{r^{3/4}}/n$$

$$\leq (4r^{11/8} + 2r^{3/4})/r^{7/5} \leq 6/r^{1/40} \leq 1/10m$$

for $r_0 \geq (60m)^{40}$, and similarly for $|U_{r+i}^n|$, $i = 1, \ldots, m-1$.

Therefore, $\mathbb{P}(\text{all } |U_r^{[n]}|, \ldots, |U_{r+m-1}^{[n]}| \geq r^{3/4}) \geq 9/10$.

We say a point $x$ selects a pair of sets $(X_1, X_2)$ if $D_x = \{x_1, x_2\}$ for some $x_1 \in X_1$ and $x_2 \in X_2$, that is, if $x$ selects a point from each set $X_1$ and $X_2$. Using the lower

bounds on $U_{r+i}^{[n]}$ we can show that, with high probability, there exist points in $B_2[2n]$ selecting each pair $(U_{r+i}^{[n]}, U_{r+j}^{[n]})$. We might hope for these to form the maximal points of a copy of $P(1, 2; m)$, since for each pair of minimal points $r + i, r + j$ we have a point above both. However, it is possible for these potential maximal points to be above more than 2 minimal points. We need to find points above exactly 2 of the minimal points. To do this we need to look at a subset of $U_{r+i}$, namely the set of points above $r + i$ but not above any other $r + j$ for $j \neq i$.

For points $x, y$ in $B_2$, write $U_{xy}$ for the set of points above both $x$ and $y$. Consider the restricted poset $B_2[n]$ and write $U_{xy}^{[n]}$ for the set of points in $B_2[n]$ above both $x$ and $y$. We will show that $|U_{r\,r+1}^{[n]}|$ is small in comparison to $|U_r^{[n]}|$ and $|U_{r+1}^{[n]}|$. Call a sequence of integers $(i_j)_{j=1}^s$ from $[r, n]$ a *path* if $i_j$ selects $i_{j-1}$ in the poset, for $j = 2, \ldots, s$. So a path is necessarily a strictly increasing sequence. We say a path $(i_j)_{j=1}^s$ is from $i_1$ to $i_s$. Define a *forked path* with ends $x, y, z$ and connection point $w$ to be three paths, one from $x$ and one from $y$ both to $w$, and a third from $w$ to $z$ (so $x, y < w \leq z$), with $w$ the only common point of the first two paths. Note that we allow the possibility that $w = z$, in which case the third path is the single point $w = z$.

For each point $u$ in $U_{r\,r+1}^{[n]}$ there must be paths $P_r$ from $r$ to $u$ and $P_{r+1}$ from $r+1$ to $u$; if we set $v = \min\{j : j$ is a common point of $P_r$ and $P_{r+1}\}$ then by taking the subpath (subsequence of consecutive terms of a path) from $r$ to $v$ (of $P_r$), the subpath from $r + 1$ to $v$ (of $P_{r+1}$) and the subpath from $v$ to $u$ (of either $P_r$ or $P_{r+1}$) we have a forked path with ends $r, r + 1$ and $u$, and connection point $v$. This forked path is not necessarily unique, since $P_r$ and $P_{r+1}$ are not necessarily unique. Let $FP(r, r + 1, v)$ be the total number of forked paths with ends $r$ and $r + 1$ and connection point $v$ all fixed, and with arbitrary third end $u$, with $v \leq u \leq n$. Let $FP(r, r + 1) = \sum_{v=r+2}^n FP(r, r + 1, v)$. Then $|U_{r\,r+1}| \leq FP(r, r + 1)$.

Now, the probability that a strictly increasing sequence $(i_j)_{j=1}^s$ is a path in $B_2[n]$ is $\mathbb{P}\left(\cap_{j=2}^s (i_j \text{ selects } i_{j-1})\right) = \prod_{j=2}^s (2/i_j)$, by independence.

We can also calculate the probability that the points $\{i_0, i_1, \ldots, i_s\}, i_0 < i_1 < \cdots < i_s$ form two disjoint paths in $B_2[n]$, one from $i_0$, the other from $i_1$, as follows. Start with two sequences $A = (i_0)$ and $B = (i_1)$, then taking each point $i_j, j = 2, \ldots, s$ in turn make it the next term in either sequence $A$ or sequence $B$. (So, the resulting $A$ and $B$ are disjoint subsequences of $(i_j)_{j=0}^s$). The probability that we can make $A$ and $B$ paths is the probability that at each step $i_j$ selects one of the current end terms of $A$ or $B$. For step $j$ this is at most $4/i_j$ so by independence the total probability is less than $\prod_{j=2}^s (4/i_j)$. We have inequality here because we are over-counting the case where $i_j$ is above both of the current end terms of $A$ and $B$.

The expected size of $FP(r, r+1, v)$ is the sum over all subsets $I$ of $[r, n]$, with $r, r+1, v \in I$, of the probability that $I$ forms a forked path with ends $r, r+1, \max I$ and connection point $v$. This is the probability that $I_{<v} = \{i \in I : i < v\}$ forms two disjoint paths from $r$ and $r+1$; and $v$ selects the end of both paths; and $I_{\geq v} = \{i \in I : i \geq v\}$ forms a path from $v$ to $\max I$. So, for $I = \{r, r+1, i_2, \ldots, i_{s-1}, v, i_{s+1}, \ldots, i_{s+s'}\}$ with $i_j$ increasing and $r+1 < i_2$, $i_{s-1} < v < i_{s+1}$ this probability is less than $\prod_{j=2}^{s-1} (4/i_j) \times 1/\binom{v}{2} \times \prod_{j=1}^{s'} (2/i_{s+j})$.

So the sum over all such subsets $I$ can be written as the following product, since the individual terms of the expanded product correspond exactly to the required probabilities for all subsets $I$,

$$\mathbb{E}FP(r, r+1, v) \leq \prod_{i=r+2}^{v-1} \left(1 + \frac{4}{i}\right) \frac{1}{\binom{v}{2}} \prod_{i=v+1}^{n} \left(1 + \frac{2}{i}\right)$$

$$\leq \exp\left\{4 \sum_{i=r+2}^{v-1} \frac{1}{i}\right\} \frac{2}{v(v-1)} \exp\left\{2 \sum_{i=v+1}^{n} \frac{1}{i}\right\}$$

$$\leq \left(\frac{v-1}{r+1}\right)^4 \frac{2}{v(v-1)} \left(\frac{n}{v}\right)^2$$

$$\leq \frac{2n^2}{r^4},$$

using the inequalities $1 + x \leq e^x$ and $\sum_{i=a}^b f(i) \leq \int_{a-1}^b f(x)dx$ for $f$ decreasing, so in particular $\sum_{i=a}^b 1/i \leq \log b - \log(a-1)$.

Therefore, $\mathbb{E}FP(r, r+1) < 2n^3/r^4$ and since $n = r^{7/5}$, we have $\mathbb{E}|U_{r,r+1}^{[n]}| <$

$\mathbb{E}FP(r, r + 1) \leq 2r^{1/5}$. The same method gives the same upper bound on the expected size of $U_{xy}^{[n]}$ for all pairs $(x, y)$ in $[r, r+m-1]^{(2)}$ so $\mathbb{P}(|U_{rr+1}^{[n]}| \geq (10m^2)r^{1/5}) \leq 1/5m^2$ and $\mathbb{P}(\text{all } |U_{xy}^{[n]}| \leq (10m^2)r^{1/5}) \geq 9/10$.

Let $A_r^{[n]}$ be the set of points above $r$ but not above $r + 1, \ldots, r + m - 1$ in $B_2[n]$, then $A_r^{[n]} = U_r^{[n]} \setminus \bigcup_{i=1}^{m-1} U_{rr+i}^{[n]}$. Similarly define $A_x^{[n]}$, $x \in [r + 1, r + m - 1]$. Then, for $r \geq r_0 \geq 400m^6$, we have $(10m^2)r^{1/5} < r^{3/4}/2m$ so with probability greater than $4/5$ we have all $|A_x^{[n]}|$, $x \in [r, r + m - 1]$ at least $\frac{1}{2}r^{3/4}$.

We grow the poset by adding a further $n = r^{7/5}$ points, to find our maximal points: $M = \binom{m}{2}$ points $a_1, \ldots, a_M$, so that each pair of sets $(A_x^{[n]}, A_y^{[n]})$, $(x, y) \in [r, r + m - 1]^{(2)}$ is selected by some $a_i$.

Now,

$$\mathbb{P}\big(n + i \text{ selects } (A_r^{[n]}, A_{r+1}^{[n]})\big) = \frac{|A_r^{[n]}||A_{r+1}^{[n]}|}{\binom{n+i}{2}} \geq \frac{r^{3/2}}{2(n + i)^2} \geq \frac{r^{3/2}}{8n^2} \quad \text{for } i \leq n,$$

so

$$\mathbb{P}\big(\text{none of } n + 1, \ldots, 2n \text{ selects } (A_r^{[n]}, A_{r+1}^{[n]})\big) \leq \left(1 - \frac{r^{3/2}}{8n^2}\right)^n$$

$$\leq \exp\left\{\frac{-r^{3/2}}{8n}\right\}$$

$$\leq \exp(-r^{1/10}/8),$$

which is less than $1/10M$ for $r_0 \geq (8 \log 10M)^{10}$. The same calculations give the same upper bound on the probability of failing to find a point in $[n + 1, 2n]$ which selects $(A_x^{[n]}, A_y^{[n]})$ for each $(x, y) \in [r, r + m - 1]^{(2)}$, so the probability of failing to find points $a_1, \ldots, a_M$ in $[n + 1, 2n]$ as desired is less than $1/10$.

So with probability at least $3/5$ we have sets $\{r, r + 1, \ldots, r + m - 1\}$ and $\{a_1, a_2, \ldots, a_M\}$ satisfying the conditions of Lemma 2.1. Therefore $\{r, r + 1, \ldots, r + m - 1, a_1, a_2, \ldots, a_M\}$ is a copy of $P(1, 2; m)$ in $B_2[r, 2n]$. $\square$

**Theorem 2.3.** *For every $m$ there exists a copy of $P(1, 2; m)$ in $B_2$, almost surely.*

**Proof.** This follows from Proposition 2.2. Fix $m$. Let $r_0$ be given by Proposition 2.2.

To find a copy of $P(1,2;m)$ in $B_2$ we split $B_2$ into disjoint sets of the form $B_2[n_1, n_2]$ as follows.

For $i = 1, 2, \ldots$, let $r_i = 2r_{i-1}^{7/5} + 1$. By Proposition 2.2 the probability of there not being a copy of $P(1,2;m)$ in $B_2[r_i, 2r_i^{7/5}]$ is less than $2/5$, for each $i$. The probability of not finding a copy of $P(1,2;m)$ in the infinite poset $B_2$ is less than the probability of not finding a copy of $P(1,2;m)$ in every poset $B_2[r_i, 2r_i^{7/5}]$. But the sets $B_2[r_i, 2r_i^{7/5}]$ are disjoint, so the events "not finding a copy of $P(1,2;m)$ in $B_2[r_i, 2r_i^{7/5}]$" are independent. Therefore the probability of not finding a copy of $P(1,2;m)$ in the infinite poset $B_2$ is zero, as required. $\qquad\square$

**Corollary 2.4.** $B_2$ *has infinite dimension, almost surely.*

**Proof.** This is immediate, since $\dim P(1,2;m) \geq \log_2 \log_2 m$. $\qquad\square$

This tells us that, almost surely, there is no finite $d$ such that $B_2$ can be embedded into $\mathbb{R}^d$, the $d$-dimensional Euclidean space with ordering $(x_1, \ldots, x_d) < (y_1, \ldots, y_d)$ in $\mathbb{R}^d$ if $x_i < y_i$ in $\mathbb{R}$, for all $i = 1, \ldots, d$, as defined earlier. What can be said for embeddings into other partial orders? Since classical sequential growth models have been proposed as possible models of discrete space-time it would be interesting to know whether the partial orders they produce can be embedded into a $d$-dimensional Minkowski space for some finite $d$.

The Minkowski space $M^d$ is defined as the partial order on $\mathbb{R}^d$ with ordering $(x_0, \ldots, x_{d-1}) < (y_0, \ldots, y_{d-1})$ in $M^d$ if $y_0 - x_0 > \sqrt{\sum_{i=1}^{d-1}(y_i - x_i)^2}$ in $\mathbb{R}$. The Minkowski dimension of a partial order $P$ is the smallest $d$ such that $P$ can be embedded into $M^d$. It is known that a finite partial order $P$ can be embedded into $M^{d+1}$ if and only if $P$ can be represented as a $d$-sphere order. A $d$-sphere order is a partial order on a ground set of spheres in $\mathbb{R}^d$, with the ordering on the spheres given by (geometric) containment. For example, the partial order $P(1,2;m)$ can always be represented as a 3-sphere order. This is a specific case of a result of Scheinerman [26]. This means that the Minkowski dimension of $P(1,2;m)$ is at

most 4, for all $m$.

We believe that the random binary order $B_2$ has infinite Minkowski dimension. A proof of this result could follow the proof strategy of Theorem 2.3; find a family of partial orders with arbitrarily large Minkowski dimension that are almost surely contained in $B_2$. Unfortunately, the partial orders known to have large Minkowski dimension are all significantly more complex than $P(1, 2; m)$. Given the complexity of the proof of Proposition 2.2 it would be ambitious to attempt a proof using this strategy. Instead, we make the following conjecture.

**Conjecture 2.5.** *$B_2$ has infinite Minkowski dimension, almost surely.*

We justify the conjecture as follows. If the poset $B_2$ has finite Minkowski dimension, then it can be embedded into $M^d$ for some $d$. Since the model $B_2$ produces the poset $B_2$ sequentially, this means that at each stage $n$ the finite poset $B_2[n]$ can be embedded into $M^d$. However, this seems unlikely since at each stage the element $n$ selects two existing elements at random, each pair of elements being equally likely with no regard to the existing structure of the embedding of $B_2[n-1]$ in $M^d$. It seems more likely that the random nature of the model $B_2$ is such that, for large enough $n$, the poset $B_2[n]$ produced at stage $n$ cannot be embedded into $M^d$.

## 2.2 Up-sets of vertices in $B_2$

Brightwell [11] proved that, almost surely, each element of $B_2$ is comparable with all but finitely many others. This result is contained within what we prove here; we need a more refined version, providing an estimate of the number of elements in $B_2[n]$ that are incomparable with an element $r$, and an estimate of the largest element incomparable with $r$. Recall that $U_r^{[n]}$ is the up-set of $r$ in $B_2[n]$ and that $I_r^{[n]} = [r, n] \setminus U_r^{[n]}$ is the set of points larger than $r$ and incomparable with $r$. We study the size $|U_r^{[n]}|$ and give good estimates of how $|U_r^{[n]}|$ grows with $n$. We then

use these estimates to provide estimates of the size $|I_r^{[n]}|$.

In [32, 33], Wormald presented a theorem which describes when and how a discrete time Markov process can be approximated by the solution to a related differential equation. However the approximation is only in terms of asymptotic bounds; here we state and prove a version of the theorem which gives explicit expressions for the approximation.

We begin with some definitions.

**Definition 2.6.** A function $f : \mathbb{R}^2 \to \mathbb{R}$ satisfies a *Lipschitz condition* on a connected open set $\mathcal{D} \subseteq \mathbb{R}^2$ if there exists a constant $L > 0$ with the property

$$|f(x_1, y_1) - f(x_2, y_2)| \leq L(|x_1 - x_2| + |y_1 - y_2|) \tag{2.7}$$

for all $(x_1, y_1)$ and $(x_2, y_2)$ in $\mathcal{D}$.

**Definition 2.7.** For $Y$ a real variable of a discrete time random process $G_0, G_1, \ldots$ which depends on a scale parameter $n$, we write $Y(t)$ for $Y(G_t)$, and for a connected set $\mathcal{D} \subseteq \mathbb{R}^2$ define the *stopping time* $T_\mathcal{D} = T_\mathcal{D}(Y)$ to be the minimum $t$ such that $(t/n, Y(t)/n) \notin \mathcal{D}$.

**Definition 2.8.** A sequence of random variables $Y_0, Y_1, \ldots$ is a *martingale* with respect to a sequence of $\sigma$-algebras $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots$ if, for all $i$,

(i) $Y_i$ is $\mathcal{F}_i$-measurable,

(ii) $\mathbb{E}|Y_i| < \infty$,

(iii) $\mathbb{E}(Y_{i+1} \mid \mathcal{F}_i) = Y_i$ almost surely.

If, instead of (iii), we have:

- $\mathbb{E}(Y_{i+1} \mid \mathcal{F}_i) \leq Y_i$ almost surely, then $(Y_i)$ is a *supermartingale* with respect to $(\mathcal{F}_i)$,

- $\mathbb{E}(Y_{i+1} \mid \mathcal{F}_i) \geq Y_i$ almost surely, then $(Y_i)$ is a *submartingale* with respect to $(\mathcal{F}_i)$.

The following lemma will be used in the theorem and is a simple extension of a martingale inequality, known as Azuma's inequality [5], to supermartingales. We omit the proof, which can be obtained by an obvious modification to the proof of Azuma's inequality.

**Lemma 2.9.** *Let $Y_0, Y_1, \ldots$ be a supermartingale with respect to a sequence of $\sigma$-algebras $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots$ with $\mathcal{F}_0$ trivial, and suppose $Y_0 = 0$ and $|Y_{i+1} - Y_i| \leq c$ for $i \geq 0$ always. Then for all $\alpha > 0$,*

$$\mathbb{P}(Y_i \geq \alpha c) \leq \exp\left(-\alpha^2/2i\right).$$

We are now in a position to state and prove our version of the theorem.

**Theorem 2.10.** *Let $Y$ be a real-valued function of the components of a discrete time Markov process $\{G_t\}_{t \geq 0}$. Assume that $\mathcal{D} \subseteq \mathbb{R}^2$ is connected, closed and bounded and contains the set*

$$\{(0, y) : \mathbb{P}(Y(0) = yn) \neq 0 \text{ for some non-negative integer } n\}$$

*and*

*(i) for some constant $\beta$,*

$$|Y(t+1) - Y(t)| \leq \beta$$

*always for $t < T_{\mathcal{D}}$,*

*(ii) for some function $f : \mathbb{R}^2 \to \mathbb{R}$ which is Lipschitz with constant $L$ on some bounded connected open set $\mathcal{D}_0$ containing $\mathcal{D}$, and some constant $\lambda$,*

$$|\mathbb{E}(Y(t+1) - Y(t)|G_t) - f(t/n, Y(t)/n)| \leq \lambda/n$$

*for $t < T_{\mathcal{D}}$,*

*(iii) $f : \mathbb{R}^2 \to \mathbb{R}$ is bounded on $\mathcal{D}_0$, i.e., there is a constant $\gamma$ such that $|f(x, y)| \leq \gamma$ for all $(x, y) \in \mathcal{D}_0$.*

*Let $w = w(n)$ be a fixed integer-valued function with $w = o(n)$. Then the following are true.*

*(a) For $(0, \hat{y}) \in \mathcal{D}$ the differential equation*

$$\frac{dy}{dx} = f(x, y)$$

*has a unique solution $y = y(x)$ in $\mathcal{D}$ passing through $y(0) = \hat{y}$, and which extends for some positive $x$ past some point, at which $x = \sigma$ say, at the boundary of $\mathcal{D}$;*

*(b) Writing $i_0 = \min\{\lfloor T_{\mathcal{D}}/w \rfloor, \lfloor \sigma n/w \rfloor\}$ and $k_i = iw$, there exists some $B > 0$ such that*

$$\mathbb{P}\big(|Y(t) - ny(t/n)| \geq B_i + (\beta + \gamma)w\big) \leq 2ie^{-2w^3/n^2}$$

*for all $i = 0, 1, \ldots, i_0 - 1$ and all $t$, $k_i \leq t \leq k_{i+1}$, and for $i = i_0$ and $k_{i_0} \leq t \leq \min\{T_{\mathcal{D}}, \sigma n\}$, where $B_i = \big((1 + Lw/n)^i - 1\big)Bw/L$, and $y(x)$ and $\sigma$ are as in (a) with $\hat{y} = Y(0)/n$.*

**Proof.** Following the proof in [32], we have part (a) from the theory of differential equations. Let $y(x)$ and $\sigma$ be as in part (a).

Let $0 \leq t \leq T_{\mathcal{D}} - w$ and let $0 \leq k < w$. This implies that $t + k < T_{\mathcal{D}}$ and so $\left(\frac{t+k}{n}, \frac{Y(t+k)}{n}\right) \in \mathcal{D}$.

By (i), we have $|Y(t + k + 1) - Y(t + k)| \leq \beta$. Also, by (ii),

$$\begin{aligned}
\mathbb{E}\big(Y(t+k+1) - Y(t+k)|G_{t+k}\big) &\leq f\left(\frac{t+k}{n}, \frac{Y(t+k)}{n}\right) + \frac{\lambda}{n} \\
&\leq f\left(\frac{t}{n}, \frac{Y(t)}{n}\right) + L\left(\frac{k}{n} + \frac{|Y(t+k) - Y(t)|}{n}\right) + \frac{\lambda}{n} \\
&\leq f\left(\frac{t}{n}, \frac{Y(t)}{n}\right) + \frac{L(w + \beta w) + \lambda}{n},
\end{aligned}$$

where the second inequality follows from (2.7). Writing $g(n)$ for $(L(w + \beta w) + \lambda)/n$, the inequality becomes

$$\mathbb{E}(Y(t + k + 1) - Y(t + k)|G_{t+k}) \leq f\left(\frac{t}{n}, \frac{Y(t)}{n}\right) + g(n).$$

Therefore, conditional on $G_t$,

$$Y(t+k) - Y(t) - kf\left(\frac{t}{n}, \frac{Y(t)}{n}\right) - kg(n)$$

is a supermartingale in $k$ with respect to the sequence of $\sigma$-fields generated by $G_t, \ldots, G_{t+w}$. The differences of the supermartingale are, by (i) and (iii), at most

$$\beta + f\left(\frac{t}{n}, \frac{Y(t)}{n}\right) + g(n) \leq \beta + \gamma + g(n).$$

So, by Lemma 2.9, for all $\alpha > 0$,

$$\mathbb{P}\left(Y(t+w) - Y(t) - wf\left(\tfrac{t}{n}, \tfrac{Y(t)}{n}\right) - wg(n) \geq \alpha(\beta + \gamma + g(n))\right) \leq e^{-\alpha^2/2w}. \quad (2.8)$$

The same argument with

$$Y(t+k) - Y(t) - kf\left(\frac{t}{n}, \frac{Y(t)}{n}\right) + kg(n)$$

a submartingale gives

$$\mathbb{P}\left(Y(t+w) - Y(t) - wf\left(\tfrac{t}{n}, \tfrac{Y(t)}{n}\right) + wg(n) \leq -\alpha(\beta + \gamma + g(n))\right) \leq e^{-\alpha^2/2w}. \quad (2.9)$$

Setting $\alpha = 2w^2/n$ and combining (2.8) and (2.9) gives

$$\mathbb{P}\left(|Y(t+w) - Y(t) - wf\left(\tfrac{t}{n}, \tfrac{Y(t)}{n}\right)| \geq 2(w^2/n)(\beta + \gamma + g(n)) + wg(n)\right) \leq 2e^{-2w^3/n^2}. \quad (2.10)$$

Now, define $k_i = iw$, $i = 0, 1, \ldots, i_0$ where $i_0 = \min\{\lfloor T_\mathcal{D}/w \rfloor, \lfloor \sigma n/w \rfloor\}$. We show by induction that for each such $i$,

$$\mathbb{P}\left(|Y(k_i) - y(k_i/n)n| \geq B_i\right) \leq 2i e^{-2w^3/n^2} \quad (2.11)$$

where $B_i = \left((1 + Lw/n)^i - 1\right)Bw/L$ for some $B > 0$.

The induction begins by the fact that $y(0) = Y(0)/n$. (Take $\hat{y} = Y(0)/n$ and use part (a).)

So, assume (2.11) is true for $i$. Write

$$A_1 = Y(k_i) - y(k_i/n)n$$

$$A_2 = Y(k_{i+1}) - Y(k_i)$$

$$A_3 = y(k_i/n)n - y(k_{i+1}/n)n$$

The inductive hypothesis (2.11) gives $|A_1| < B_i$ with probability at least $1 - 2ie^{-2w^3/n^2}$. By (2.10) we have

$$|A_2 - wf(k_i/n, Y(k_i)/n)| < 2(w^2/n)(\beta + \gamma + g(n)) + wg(n)$$

with probability at least $1 - 2e^{-2w^3/n^2}$.

Since $f$ satisfies the Lipschitz condition and $(k_{i+1}/n, Y(k_{i+1})/n) \in \mathcal{D}$ (because $k_{i+1} < T_\mathcal{D}$), we also have

$$
\begin{aligned}
|A_3 + wy'(k_i/n)| &= |y(k_i/n)n - y(k_{i+1}/n)n + wy'(k_i/n)| \\
&= |-wy'(k/n) + wy'(k_i/n)| && \text{for some } k,\ k_i \le k \le k_{i+1} \\
&= w|f(k/n, y(k/n)) - f(k_i/n, y(k_i/n))| && \text{since } y \text{ is solution to (a)} \\
&\le wL[w/n + |y(k/n) - y(k_i/n)|] && \text{by (2.7)} \\
&\le wL[w/n + (w/n)|f(k'/n, y(k'/n))|] && \text{for some } k',\ k_i \le k' \le k \\
&\le wL[w/n + (w/n)\gamma] && \text{by (iii)} \\
&= L(1+\gamma)w^2/n
\end{aligned}
$$

where we have used the Mean Value Theorem (twice, to get lines 2 and 5). So,

$$|y'(k_i/n) - f(k_i/n, Y(k_i)/n)| = |f(k_i/n, y(k_i/n)) - f(k_i/n, Y(k_i)/n)| \le L|A_1|/n$$

and so assuming $|A_1| < B_i$, we have

$$|A_3 - (-wf(k_i/n, Y(k_i)/n))| \le \frac{L(1+\gamma)w^2}{n} + \frac{Lw}{n}|A_1| \le \frac{L(1+\gamma)w^2}{n} + \frac{Lw}{n}B_i.$$

So, we have

$$|Y(k_{i+1}) - y(k_{i+1}/n)n| = |A_1 + A_2 + A_3|$$

$$< B_i + 2(w^2/n)(\beta + \gamma + g(n)) + wg(n) + L(1+\gamma)w^2/n + B_i Lw/n$$

$$= [2(w^2/n)(\beta + \gamma + g(n)) + wg(n) + L(1+\gamma)w^2/n] + B_i(1 + Lw/n) \quad (2.12)$$

with probability at least $1 - 2(i+1)e^{-2w^3/n^2}$.

There exists $B > 0$ with

$$2(w^2/n)(\beta + \gamma + g(n)) + wg(n) + L(1+\gamma)w^2/n \leq Bw^2/n \quad (2.13)$$

for all $n$, so the term on the right hand side of inequality (2.12) can be replaced with $B_i(1 + Lw/n) + Bw^2/n$, which is exactly $B_{i+1}$. So we have (2.11) for $i + 1$.

Finally, $k_{i+1} - k_i = w$ and the variation in $Y(t)$ when $t$ changes by at most $w$ is at most $\beta w$, by (i), and as before $|y(t_1/n)n - y(t_2/n)n|$ is less than $w|f(t/n, y(t/n))|$ for some $t$, $t_1 \leq t \leq t_2$ and this is less than $\gamma w$. So

$$\mathbb{P}(|Y(t) - ny(t/n)| \geq B_i + (\beta + \gamma)w) \leq 2ie^{-2w^3/n^2}$$

for all $i = 0, 1, \ldots, i_0 - 1$ and all $t$, $k_i \leq t \leq k_{i+1}$, and for $i = i_0$ and $k_{i_0} \leq t \leq \min\{T_{\mathcal{D}}, \sigma n\}$. $\square$

We can apply Theorem 2.10 to $|U_r^{[n]}|$ as follows. We take as the Markov process the random binary growth model, and as the real-valued function the size of the up-set of a fixed vertex $r$. We then find sets $\mathcal{D}$ and $\mathcal{D}_0$, a function $f$, and constants $\beta, \lambda$ and $\gamma$ satisfying the assumptions of the theorem. We obtain the following corollary, which shows fairly precisely how $|U_r^{[m]}|$ grows as $m$ goes from some initial $n$ to $(\sigma + 1)n$, where $\sigma$ is a large constant. Over this range, $|U_r^{[n]}|/n$ grows from a small value to a value near to 1.

**Corollary 2.11.** *For fixed $r$ and any $n > r$, if $|U_r^{[n]}| = c(n)n$ for $c(n)$ an arbitrary function of $n$, then*

$$\mathbb{P}\left(\left|\frac{|U_r^{[n(\sigma+1)]}|}{n(\sigma+1)} - \frac{\sigma+1}{\sigma + 1/c(n)}\right| \geq \left(\frac{10e^{2.1\sigma} + 2.1}{\sigma + 1}\right)\frac{1}{n^{1/3-\delta}}\right) \leq 2\sigma n^{1/3-\delta}e^{-2n^{3\delta}}$$

*for any constants $0 < \delta < 1/3$, $\sigma > 0$.*

**Proof.** Fix a vertex $r$ in $B_2[n]$. Let the Markov process $\{G_t\}_{t \geq 0}$ be the random binary growth model but starting at stage $n$, so that $G_t$ corresponds to $B_2[n + t]$. Let $Y(t)$ be the size of the up-set of $r$ in $B_2[n + t]$, i.e., $Y(t) = |U_r^{[n+t]}|$. For any constant $\sigma$, define $\mathcal{D}$ as the region $\{(x, y) : 0 \leq x \leq \sigma, 0 \leq y \leq x + 1\}$. The region $\mathcal{D}$ contains the interval $\{(0, y) : 0 \leq y \leq 1\}$, and since $|U_r^{[n]}| = c(n)n$, we must have $c(n) \leq 1$ for all $n$. So, $\mathcal{D}$ satisfies the assumption in Theorem 2.10, since it contains all points $(0, c(n))$ for $n = 1, 2, \ldots$. We now find a set $\mathcal{D}_0$, a function $f$, and constants $\beta$, $\lambda$ and $\gamma$ satisfying assumptions (i)–(iii).

Since $Y(t) = |U_r^{[n+t]}| \leq n + t$ we have $Y(t)/n \leq t/n + 1$, and so $(t/n, Y(t)/n) \in \mathcal{D}$ as long as $t/n \leq \sigma$. This implies $T_{\mathcal{D}} = \lfloor \sigma n \rfloor + 1$.

Let $\beta = 1$, then (i) holds since $|Y(t+1) - Y(t)| = |U_r^{[n+t+1]}| - |U_r^{[n+t]}| \leq 1$ always for $t \leq \sigma n$.

Let $f(x, y) = 2y/(x + 1) - y^2/(x + 1)^2$. Let $L = 2.1$ and $\gamma = 1.1$. The function $f$ is bounded on $\mathcal{D}$ by 1 (attained when $y = x + 1$) and is continuous over the boundary of $\mathcal{D}$, so there exists an open set $\mathcal{D}'$ containing $\mathcal{D}$ on which $f$ is bounded by $\gamma = 1.1$. Also, $\|\nabla f\|$, the length of the gradient vector of $f$ ($\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$), is bounded on $\mathcal{D}$ by 2 and is continuous over the boundary of $\mathcal{D}$, so there exists an open set $\mathcal{D}''$ containing $\mathcal{D}$ on which $\|\nabla f\|$ is bounded by $L = 2.1$. But then

$$|f(\mathbf{u}) - f(\mathbf{v})| \leq L|\mathbf{u} - \mathbf{v}| \tag{2.14}$$

for all $\mathbf{u}, \mathbf{v} \in \mathcal{D}''$, so $f$ is Lipschitz with constant $L$ on $\mathcal{D}''$ (this follows by applying the triangle inequality to the right hand side of (2.14)). Let $\mathcal{D}_0$ be the intersection

of the two sets $\mathcal{D}', \mathcal{D}''$. So, (iii) holds, and (ii) holds with $\lambda = 1$, since

$$\mathbb{E}(Y(t+1) - Y(t)|G_t) = 0 \times \mathbb{P}(Y(t+1) = Y(t)|G_t) + 1 \times \mathbb{P}(Y(t+1) = Y(t) + 1|G_t)$$

$$= 1 - \binom{n+t+1-Y(t)}{2} / \binom{n+t+1}{2}$$

$$= 1 - \frac{(n+t+1-Y(t))(n+t-Y(t))}{(n+t+1)(n+t)}$$

$$= \frac{2Y(t)(n+t+1) - Y(t)(Y(t)+1)}{(n+t+1)(n+t)},$$

which differs from $f(t/n, Y(t)/n)$ by at most $1/n$ for $t \leq \sigma n$.

Now $T_{\mathcal{D}} = \lfloor \sigma n \rfloor + 1$ and so $T_{\mathcal{D}} > \sigma n$. So Theorem 2.10 gives the result (b) for $i = i_0$, $t = \sigma n$, namely that, for some $B > 0$,

$$\mathbb{P}(|Y(\sigma n) - ny(\sigma)| > B_{i_0} + 2.1w) \leq 2i_0 e^{-2w^3/n^2}. \tag{2.15}$$

Here $y(x)$ is the solution to the differential equation

$$\frac{dy}{dx} = 2\frac{y}{x+1} - \frac{y^2}{(x+1)^2}$$

with initial condition $y(0) = c(n)$. This is a homogeneous equation with solution

$$y(x) = \frac{(x+1)^2}{x + 1/c(n)}.$$

Also, $i_0 \leq \sigma n/w$, so $B_{i_0} = ((1 + Lw/n)^{i_0} - 1) Bw/L \leq Bwe^{L\sigma}/L$, and (2.15) becomes

$$\mathbb{P}\left(\left||U_r^{[n+\sigma n]}| - n\frac{(\sigma+1)^2}{\sigma + 1/c(n)}\right| > Bwe^{L\sigma}/L + 2.1w\right) \leq 2(\sigma n/w)e^{-2w^3/n^2}.$$

for some $B > 0$.

Choose $\delta$ with $0 < \delta < 1/3$ and set the arbitrary function $w(n)$ to $n^{2/3+\delta}$. Then $w(n) = o(n)$ and so using the particular values for $L, \beta, \gamma$ and $\lambda$, we can satisfy equation (2.13) with $B = 21$ and this gives the required result. $\qquad\square$

In the proof of Proposition 2.2 we bounded the expectation of the hitting time of the event $|U_r| = k$. We use this bound to show that $U_r$ contains all but finitely many points of $B_2$, almost surely. In terms of $I_r$ we have the following theorem.

**Theorem 2.12.** *For any constants $\varepsilon, \eta$ with $0 < \varepsilon < 1/4$ and $0 < \eta < 1$ there exists $r_0$ such that for all $r \geq r_0$ both $|I_r| < r^{2+4\varepsilon}$ and $I_r \subseteq [r, r^{4+8\varepsilon}]$ hold with probability at least $1 - \eta$.*

**Proof.** Assume that $r$ is sufficiently large. As before, let $T_k$ be the hitting time of event $|U_r| = k$, in terms of the growth model, i.e., the smallest $t$ such that $|U_r^{[t]}| = k$. As in (2.3), we have $\mathbb{E} T_k \leq 2r\sqrt{k} + 2k$. So $\mathbb{E} T_{r^2} \leq 4r^2$ and Markov's inequality gives

$$\mathbb{P}\left(|U_r^{[(16/\eta)r^2]}| < r^2\right) = \mathbb{P}\left(T_{r^2} > (16/\eta)r^2\right) < \eta/4 \tag{2.16}$$

so that with suitably high probability the size of the up-set, $|U_r^{[(16/\eta)r^2]}|$, is at least fraction $\eta/16$ of the size of the poset, $(16/\eta)r^2$.

Set $n_0 = (16/\eta)r^2$. We can rewrite equation (2.16) as

$$\mathbb{P}\left(|U_r^{[n_0]}|/n_0 \geq \eta/16\right) > 1 - \eta/4. \tag{2.17}$$

Assume we have $|U_r^{[n_0]}|/n_0 \geq \eta/16$. Let $\varepsilon$ be an arbitrary constant with $0 < \varepsilon < 1/4$. We will use Corollary 2.11 to show that as the size of the poset, $n$, increases from $n_0$ to $(\sigma + 1)n_0$, for some constant $\sigma$, the ratio $|U_r^{[n]}|/n$ also increases, to a value that is at least $1 - \varepsilon/2$.

**Claim 2.1.** *There exists a constant $\sigma_0$ (dependent on $\varepsilon$ and $\eta$) such that if $\dfrac{|U_r^{[n_0]}|}{n_0} \geq \eta/16$ then $\dfrac{|U_r^{[(\sigma_0+1)n_0]}|}{(\sigma_0 + 1)n_0} \geq 1 - \varepsilon/2$ with probability at least $2\sigma_0 n_0^{1/4} e^{-2n_0^{1/4}}$.*

**Proof of Claim 2.1.** Suppose $|U_r^{[n_0]}|/n_0 \geq \eta/16$. Applying Corollary 2.11 with $n = n_0$, $c(n_0) = \eta/16$ and $\delta = 1/12$ we have

$$\mathbb{P}\left(\left|\frac{|U_r^{[n_0(\sigma+1)]}|}{n_0(\sigma+1)} - \frac{\sigma+1}{\sigma+16/\eta}\right| \geq \left(\frac{10e^{2.1\sigma} + 2.1}{\sigma+1}\right)\frac{1}{n_0^{1/4}}\right) \leq 2\sigma n_0^{1/4} e^{-2n_0^{1/4}} \tag{2.18}$$

for any $\sigma > 0$. Set $\sigma_0$ so that

$$\frac{\sigma_0 + 1}{\sigma_0 + 16/\eta} = 1 - \varepsilon/4 \tag{2.19}$$

and then for sufficiently large $r$, $\left(\dfrac{10e^{2.1\sigma_0} + 2.1}{\sigma_0 + 1}\right)\dfrac{1}{n_0^{1/4}} \leq \varepsilon/4$. Combining this inequality with (2.18) and (2.19) and setting $\sigma = \sigma_0$ gives the result. $\qquad\square$

Let $M = (16/\eta)(\sigma_0 + 1)$, so that $(\sigma_0 + 1)n_0 = Mr^2$. We have shown that, with suitably high probability, $|U_r^{[n]}|/n \geq 1 - \varepsilon/2$ for $n = Mr^2$. We now show that $|U_r^{[n]}|/n$ remains close to 1 for all larger $n$. That is, that $|U_r^{[n]}|/n \geq 1 - \varepsilon$ for all $n \geq Mr^2$.

Let $n_1 = Mr^2$, and $n_i = (1 + \varepsilon/2)^{i-1}n_1$ for $i = 2, 3, \ldots$.

**Claim 2.2.** *If $|U_r^{[n_i]}|/n_i \geq 1 - \varepsilon/2$ then*

*(a) $|U_r^{[n]}|/n \geq 1 - \varepsilon$ for $n = n_i + 1, n_i + 2, \ldots, \lfloor n_{i+1}\rfloor$ and*

*(b) $|U_r^{[n_{i+1}]}|/n_{i+1} \geq 1 - \varepsilon/2$ with probability at least $1 - \varepsilon n_i^{1/4}e^{-2n_i^{1/4}}$.*

**Proof of Claim 2.2.** Suppose we have $|U_r^{[n_i]}|/n_i \geq 1 - \varepsilon/2$.

For part (a) we use the fact that $|U_r^{[n]}|$ is increasing in $n$, so that

$$\frac{|U_r^{[n]}|}{n} \geq \frac{|U_r^{[n_i]}|}{n_{i+1}} = \frac{|U_r^{[n_i]}|}{(1 + \varepsilon/2)n_i} \geq \frac{1 - \varepsilon/2}{1 + \varepsilon/2} \geq 1 - \varepsilon$$

for all $n = n_i + 1, n_i + 2, \ldots, \lfloor n_{i+1}\rfloor$.

For part (b) we apply Corollary 2.11 with $n = n_i$, $\sigma = \varepsilon/2$ and $\delta = 1/12$. We have

$$\mathbb{P}\left(\left|\frac{|U_r^{[n_i(\varepsilon/2+1)]}|}{n_i(\varepsilon/2 + 1)} - \frac{\varepsilon/2 + 1}{\varepsilon/2 + 1/c(n_i)}\right| \geq \left(\frac{10e^{2.1\varepsilon/2} + 2.1}{\varepsilon/2 + 1}\right)\frac{1}{n_i^{1/4}}\right) \leq \varepsilon n_i^{1/4}e^{-2n_i^{1/4}} \tag{2.20}$$

with $c(n_i) \geq 1 - \varepsilon/2$. So,

$$\frac{\varepsilon/2 + 1}{\varepsilon/2 + 1/c(n_i)} \geq \frac{\varepsilon/2 + 1}{\varepsilon/2 + 1/(1 - \varepsilon/2)}$$

and for sufficiently large $r$,

$$\frac{\varepsilon/2 + 1}{\varepsilon/2 + 1/(1 - \varepsilon/2)} - \left(\frac{10e^{2.1\varepsilon/2} + 2.1}{\varepsilon/2 + 1}\right)\frac{1}{n_i^{1/4}} \geq 1 - \varepsilon/2. \tag{2.21}$$

Then, (2.20) becomes $\mathbb{P}\left(|U_r^{[n_{i+1}]}|/n_{i+1} \leq 1 - \varepsilon/2\right) \leq \varepsilon n_i^{1/4} e^{-2n_i^{1/4}}$.     $\square$

Notice that, since $n_{i+1} > n_i$, if the inequality (2.21) is satisfied for $i = 1$, then it is automatically satisfied for all larger $i$. That is, if we have $r$ sufficiently large to be able to apply Claim 2.2 once, then we can apply it repeatedly to get the following.

Assuming $|U_r^{[n_1]}|/n_1 \geq 1 - \varepsilon/2$, we have $|U_r^{[n]}|/n \geq 1 - \varepsilon$ for all integers $n \geq n_1 = Mr^2$ with probability at least $1 - \sum_{i=1}^{\infty} \varepsilon n_i^{1/4} e^{-2n_i^{1/4}}$, for sufficiently large $r$.

Let $r$ be sufficiently large so that $2\sigma_0 n_0^{1/4} e^{-2n_0^{1/4}} + \sum_{i=1}^{\infty} \varepsilon n_i^{1/4} e^{-2n_i^{1/4}} < \eta/4$. Then, we have $|U_r^{[n]}|/n \geq 1 - \varepsilon$ for all integers $n \geq Mr^2$ with probability at least $1 - \eta/2$. Once $|U_r^{[n]}|$ is always a large fraction of $n$, we can show that $U_r^{[n]}$ becomes almost all of the poset $B_2[n]$ for $n = r^{4+8\varepsilon}$. Rather, we now look at $I_r^{[n]}$, the set of points in $[r, n]$ incomparable with $r$ in $B_2[n]$.

For $t \geq Mr^2$, set $s_t = |I_r^{[t]}|/\sqrt{t}$, and consider the sequence $(s_t)$ as a stochastic process.

We have that

$$s_{t+1} = \begin{cases} s_t \dfrac{\sqrt{t}}{\sqrt{t+1}} & \text{with probability } 1 - \binom{|I_r^{[t]}|}{2}\big/\binom{t+1}{2} \\[2ex] \dfrac{s_t\sqrt{t}+1}{\sqrt{t+1}} & \text{with probability } \binom{|I_r^{[t]}|}{2}\big/\binom{t+1}{2} \end{cases}$$

Therefore

$$\mathbb{E}s_{t+1} = \frac{s_t\sqrt{t} + \binom{s_t\sqrt{t}}{2}\big/\binom{t+1}{2}}{\sqrt{t+1}} = s_t\sqrt{\frac{t}{t+1}}\left(1 + \frac{s_t\sqrt{t}-1}{t(t+1)}\right).$$

Now, provided $s_t \leq \varepsilon\sqrt{t}$ (which will be the case unless $|U_r^{[t]}|$ drops below $(1-\varepsilon)t$), we have

$$\mathbb{E}s_{t+1} \leq s_t\left(1 - \frac{1}{t+1}\right)^{1/2}\left(1 + \frac{\varepsilon}{t+1}\right) \leq s_t\left(1 - \frac{1/2 - \varepsilon}{t+1}\right)$$

for all $t \geq Mr^2$. So

$$\mathbb{E}s_{Mr^2+k} \leq s_{Mr^2} \prod_{j=1}^{k} \left(1 - \frac{1/2 - \varepsilon}{Mr^2 + j}\right) \leq s_{Mr^2} \exp\left(-(1/2 - \varepsilon)\sum_{j=1}^{k} \frac{1}{Mr^2 + j}\right)$$

$$\leq \varepsilon\sqrt{Mr^2}\left(\frac{Mr^2 + 1}{Mr^2 + k + 1}\right)^{1/2-\varepsilon}$$

$$\leq \frac{\sqrt{Mr^2}}{4}\left(\frac{Mr^2 + 1}{Mr^2 + k + 1}\right)^{1/2-\varepsilon},$$

where we have used the fact that $\varepsilon < 1/4$ to get the last line. So,

$$\mathbb{E}s_{r^{4+8\varepsilon}} \leq \sqrt{Mr^2}\left(\frac{Mr^2}{r^{4+8\varepsilon}}\right)^{1/2-\varepsilon} \leq M^{1-\varepsilon}r^{-2\varepsilon+8\varepsilon^2}.$$

Using Markov's inequality, we have $s_{r^{4+8\varepsilon}} \leq (4/\eta)M^{1-\varepsilon}r^{-2\varepsilon+8\varepsilon^2}$ with probability at least $1 - \eta/4$.

Therefore $|I_r^{[r^{4+8\varepsilon}]}| \leq \sqrt{r^{4+8\varepsilon}}.(4/\eta)M^{1-\varepsilon}r^{-2\varepsilon+8\varepsilon^2} = \widetilde{M}r^{2+2\varepsilon+8\varepsilon^2}$ with probability at least $1 - \eta/4$, where $\widetilde{M} = (4/\eta)M^{1-\varepsilon}$.

Finally, let us consider the probability that all vertices with a label higher than $r^{4+8\varepsilon}$ are comparable with $r$; in other words $I_r^{[s]} = I_r^{[r^{4+8\varepsilon}]}$ for $s \geq r^{4+8\varepsilon}$. Given the size $|I_r^{[r^{4+8\varepsilon}]}|$, this probability is exactly

$$\prod_{s=r^{4+8\varepsilon}+1}^{\infty}\left(1 - \frac{\binom{|I_r^{[r^{4+8\varepsilon}]}|}{2}}{\binom{s}{2}}\right),$$

which is at least

$$1 - \frac{|I_r^{[r^{4+8\varepsilon}]}|^2}{2}\sum_{s=r^{4+8\varepsilon}+1}^{\infty}\frac{1}{\binom{s}{2}} = 1 - \frac{|I_r^{[r^{4+8\varepsilon}]}|^2}{r^{4+8\varepsilon}} \geq 1 - \frac{\widetilde{M}^2}{r^{4\varepsilon-16\varepsilon^2}}.$$

Since $\varepsilon < 1/4$ we have $4\varepsilon - 16\varepsilon^2 > 0$ so that for sufficiently large $r$, $\widetilde{M}^2/r^{4\varepsilon-16\varepsilon^2} < \eta/4$. Also, $|I_r^{[r^{4+8\varepsilon}]}| \leq \widetilde{M}r^{2+2\varepsilon+8\varepsilon^2} \leq r^{2+4\varepsilon}$, for sufficiently large $r$. So, combining all the probabilities, we have $|I_r| = |I_r^{[r^{4+8\varepsilon}]}| \leq r^{2+4\varepsilon}$ and $I_r \subseteq [r, r^{4+8\varepsilon}]$ with probability at least $1 - \eta$, as required. $\qquad\square$

This result is close to the best possible; as the following lemma shows, we have that $\mathbb{E}|U_r^{[n]}| \leq n^2/r^2$, so for small $\varepsilon > 0$, $|I_r| \geq r^{2-\varepsilon}$ with high probability.

**Lemma 2.13.** *For all $n > r$, $\mathbb{E}|U_r^{[n]}| \leq n^2/r^2$.*

**Proof.** Firstly, we make an observation similar to that in the proof of Proposition 2.2 on page 26, that for all $u \in U_r^{[n]}$ there must exist a path from $r$ to $u$. Therefore, it is enough to provide an upper bound on the expected number of paths in $B_2[n]$ with start point $r$. As before, for $r < i_1 < i_2 < \cdots < i_s$, the probability that $\{r, i_1, i_2, \ldots, i_s\}$ is a path is

$$\mathbb{P}(i_1 \text{ selects } r) \prod_{j=2}^{s} \mathbb{P}(i_j \text{ selects } i_{j-1}) = \prod_{j=1}^{s} \frac{2}{i_j}.$$

So, the expected number of paths in $B_2[n]$ starting at $r$ is bounded above by

$$\sum_{I \subseteq [r+1,n]} \mathbb{P}(\{r\} \cup I \text{ is a path}) = \sum_{I \subseteq [r+1,n]} \prod_{i \in I} \frac{2}{i} = \prod_{i=r+1}^{n} \left(1 + \frac{2}{i}\right)$$

$$= \frac{(n+1)(n+2)}{(r+1)(r+2)} \leq \frac{n^2}{r^2}. \qquad \square$$

We have shown that for a typical $r$, the size $|U_r^{[n]}|$ is a constant fraction of $n$ for $n = \Theta(r^2)$, and that the set $I_r$ is contained in $[r^{4+8\varepsilon}]$, with $|I_r| = O(r^{2+4\varepsilon})$. What about for a worst case $r$? Can we say something about all but finitely many $r$?

Clearly, we cannot always expect $|U_r^{[n]}|$ to be a constant fraction of $n$ for $n = \Theta(r^2)$. As we showed in Section 1,

$$\mathbb{P}(r \text{ is maximal in } B_2[n]) = \frac{r(r-1)}{n(n-1)}$$

which is approximately $r^2/n^2$. Setting $n = r^{3/2}$, we have that

$$\mathbb{P}(r \text{ is maximal in } B_2[r^{3/2}]) \approx \frac{1}{r}$$

which means there are infinitely many $r$ with $|U_r^{[r^{3/2}]}| = 1$. When this is the case, the growth process of $U_r^{[n]}$ for $n > r^{3/2}$ is identical to the growth process of $U_{r^{3/2}}^{[n]}$ for $n > r^{3/2}$, since the sizes of $U_r^{[r^{3/2}]}$ and $U_{r^{3/2}}^{[r^{3/2}]}$ are the same. So, the expected size of $U_r^{[n]}$ can be found by substituting $r^{3/2}$ for $r$ in Lemma 2.13, which shows that $\mathbb{E}|U_r^{[n]}| \leq n^2/(r^{3/2})^2 = n^2/r^3$. So, for such an $r$ the expected size of $U_r^{[r^2]}$ is less than $r$, and we need $n = \Theta(r^3)$ before the expected size of $U_r^{[n]}$ is a constant fraction of $n$. We believe this is the worst case, that $|U_r^{[n]}|$ is a constant fraction of $n$ for $n = \Theta(r^3)$,

and then $I_r$ is contained in $[r^{6+\varepsilon'}]$, with $|I_r| = O(r^{3+\varepsilon})$. Heuristically, it appears that the growth of $|U_r^{[n]}|$ is highly dependent on the values of the hitting times, $T_k$, for small $k$, which are not concentrated near the mean values; for example, the above argument shows that $T_2$ can be as large as $r^{3/2}$, whereas the mean $\mathbb{E}T_2$ is bounded above by $2r + 2$, using equation (2.2). Indeed, once $|U_r^{[n]}|/n$ is at least $1/n^{1/3}$ we can apply Corollary 2.11, to closely approximate the growth. However, it appears rather difficult to prove these statements in full, and we settle for the following polynomial bounds on the size $|I_r|$ and the value of the largest $s$ incomparable with $r$.

**Theorem 2.14.** *For all but finitely many $r$, $|I_r| \leq r^{27/5}$ and $I_r \subseteq [r^{12}]$.*

The proof is naturally very similar to the proof of Theorem 2.12.

**Proof.** Fix $r$. As before, let $T_k$ be the hitting time of event $|U_r| = k$, in terms of the growth model, i.e., the smallest $t$ such that $|U_r^{[t]}| = k$. As (2.3), we have $\mathbb{E}T_k \leq 2r\sqrt{k} + 2k$. So $\mathbb{E}T_{r^{13/6}} \leq 4r^{13/6}$. Markov's inequality gives

$$\mathbb{P}\left(|U_r^{[r^{3+8/45}]}| < r^{13/6}\right) = \mathbb{P}\left(T_{r^{13/6}} > r^{3+8/45}\right) < 4/r^{91/90}. \tag{2.22}$$

Set $n_0 = r^{3+8/45}$. Equation (2.22) becomes

$$\mathbb{P}\left(|U_r^{[n_0]}|/n_0 \geq 1/n_0^{7/22}\right) > 1 - 4/r^{91/90}.$$

Assume we have $|U_r^{[n_0]}|/n_0 \geq 1/n_0^{7/22}$. We will use Corollary 2.11 to show that as we increase the size of the poset by a factor of 2, the fraction $|U_r^{[n]}|/n$ also increases by a factor that is only slightly smaller than 2. We can use this method repeatedly until $|U_r^{[n]}|/n$ is at least some constant fraction.

Let $n_i = 2^i n_0$ for $i = 1, 2, \ldots$ and let $c(n) = |U_r^{[n]}|/n$ for all $n \geq n_0$.

**Claim 2.3.** *If $1/n_0^{7/22} < c(n_i) < 1/300$ then $c(n_{i+1}) \geq (149/75)c(n_i)$ with probability at least $1 - 2n_i^{8/25} e^{-2n_i^{1/25}}$.*

**Proof of Claim 2.3.** Suppose $1/n_0^{7/22} < c(n_i) < 1/300$. The upper bound on $c(n_i)$ implies

$$\frac{2}{1+1/c(n_i)} > (299/150)c(n_i) \qquad (2.23)$$

and the lower bound implies

$$\left(\frac{10e^{2.1}+2.1}{2}\right)\frac{1}{n_i^{8/25}} < (1/150)\frac{1}{n_0^{7/22}} < (1/150)c(n_i). \qquad (2.24)$$

So applying Corollary 2.11 with $n = n_i, \delta = 1/75, \sigma = 1$, we have

$$\mathbb{P}\left(\left|\frac{|U_r^{[2n_i]}|}{2n_i} - \frac{2}{1+1/c(n_i)}\right| \geq \left(\frac{10e^{2.1}+2.1}{2}\right)\frac{1}{n_i^{8/25}}\right) \leq 2n_i^{8/25}e^{-2n_i^{1/25}}$$

which, using (2.23) and (2.24), gives the result. $\qquad \square$

Using Claim 2.3 repeatedly we have that for $k = 0, 1, \ldots$ either $c(n_l) \geq 1/300$ for some $l < k$, or

$$c(n_k) \geq (149/75)^k c(n_0) \geq (149/75)^k/n_0^{7/22}$$

with probability at least $1 - \sum_{i=0}^{k-1} 2n_i^{8/25}e^{-2n_i^{1/25}}$.

So, there exists a $k \leq \dfrac{\log\left((1/300)n_0^{7/22}\right)}{\log(149/75)}$ such that $|U_r^{[n_k]}|/n_k \geq 1/300$ with probability at least $1 - (\log n_0)n_0^{1/2}e^{-2n_0^{1/25}}$.

We have

$$n_k \leq 2^{\log(n_0^{7/22}/300)/\log(149/75)}n_0 = (n_0^{7/22}/300)^{\log 2/\log(149/75)}n_0. \qquad (2.25)$$

Using $n_0 = r^{3+8/45}$ we get $n_k \leq r^{21/5}/317$.

Assume we have $|U_r^{[n_k]}|/n_k \geq 1/300$. We will apply Corollary 2.11 once more to increase the fraction $|U_r^{[n]}|/n$ to a constant close to 1.

**Claim 2.4.** $\dfrac{|U_r^{[n]}|}{n} \geq 77/78$ *with probability at least* $1 - 10^5 n_k^{1/4}e^{-2n_k^{1/4}}$, *where* $n = 46345n_k \leq 150r^{21/5}$.

**Proof of Claim 2.4.** We have $|U_r^{[n_k]}|/n_k \geq 1/300$. Applying Corollary 2.11, with $n = n_k$ and $\delta = 1/12$ we have

$$\mathbb{P}\left(\left|\frac{|U_r^{[n_k(\sigma+1)]}|}{n_k(\sigma+1)} - \frac{\sigma+1}{\sigma+1/c(n_k)}\right| \geq \left(\frac{10e^{2.1\sigma}+2.1}{\sigma+1}\right)\frac{1}{n_k^{1/4}}\right) \leq 2\sigma n_k^{1/4}e^{-2n_k^{1/4}} \qquad (2.26)$$

for any $\sigma > 0$. Set $\sigma = 46344$ so that

$$\frac{\sigma + 1}{\sigma + 1/c(n_k)} = \frac{46345}{46344 + 1/c(n_k)} \geq 155/156, \tag{2.27}$$

which is possible, since $c(n_k) \geq 1/300$. Then for sufficiently large $r$,

$$\frac{10e^{2.1\sigma} + 2.1}{\sigma + 1} \frac{1}{n_k^{1/4}} \leq 1/156.$$

Combining with (2.26) and (2.27) and setting $\sigma = 46344$ gives the result. $\qquad\square$

By a similar method we can show that $\frac{|U_r^{[t]}|}{t} \geq 77/78$ for all $t \geq n$ with probability at least $1 - \sum_{t=n}^{\infty} t^{1/4} e^{-2t^{1/4}}$.

As before, for $t \geq n$, set $s_t = |I_r^{[t]}|/\sqrt{t}$, and consider the sequence $(s_t)$ as a stochastic process. Again, we have

$$\mathbb{E}s_{t+1} = \frac{s_t\sqrt{t} + \binom{s_t\sqrt{t}}{2}/\binom{t+1}{2}}{\sqrt{t+1}} = s_t\sqrt{\frac{t}{t+1}}\left(1 + \frac{s_t\sqrt{t} - 1}{t(t+1)}\right).$$

Now, provided $s_t \leq \sqrt{t}/78$ (which will be the case unless $|U_r^{[t]}|$ drops below $(77/78)t$), we have

$$\mathbb{E}s_{t+1} \leq s_t\left(1 - \frac{1}{t+1}\right)^{1/2}\left(1 + \frac{1}{78(t+1)}\right) \leq s_t\left(1 - \frac{1/2 - 1/78}{(t+1)}\right)$$

which gives

$$\mathbb{E}s_{t+k} \leq s_t\prod_{j=1}^{k}\left(1 - \frac{19}{39(t+j)}\right) \leq s_t\exp\left(-(19/39)\sum_{j=1}^{k}\frac{1}{t+j}\right)$$
$$\leq \frac{\sqrt{t}}{78}\left(\frac{t+1}{t+k+1}\right)^{19/39}.$$

So, for example, $\mathbb{E}s_{t^{20/7}} \leq 1/(78t^{17/42})$, and for $t = r^{21/5}$ this gives $\mathbb{E}s_{r^{12}} \leq 1/r^{17/10}$. By Markov's inequality, we have $s_{r^{12}} \leq 1/r^{3/5}$ with probability at least $1 - 1/r^{11/10}$.

Therefore $|I_r^{[r^{12}]}| \leq \sqrt{r^{12}}/r^{3/5} = r^{27/5}$ with probability at least $1 - 1/r^{11/10}$.

Finally, let us consider the probability that all vertices with a label higher than $r^{12}$ are comparable with $r$; in other words $I_r^{[s]} = I_r^{[r^{12}]}$ for $s \geq r^{12}$. Given the size $|I_r^{[r^{12}]}|$, this probability is exactly

$$\prod_{s=r^{12}+1}^{\infty}\left(1 - \frac{\binom{|I_r^{[r^{12}]}|}{2}}{\binom{s}{2}}\right),$$

which is at least

$$1 - \frac{|I_r^{[r^{12}]}|^2}{2} \sum_{s=r^{12}+1}^{\infty} \frac{1}{\binom{s}{2}} = 1 - \frac{|I_r^{[r^{12}]}|^2}{r^{12}} \geq 1 - \frac{1}{r^{6/5}}.$$

So, combining all the probabilities, we have $|I_r| = |I_r^{[r^{12}]}| \leq r^{27/5}$ and $I_r^{[s]} = I_r^{[r^{12}]}$ for $s \geq r^{12}$ with probability at least

$$1 - 4/r^{91/90} - (\log n_0)n_0^{1/2}e^{-2n_0^{1/25}} - 10^5 n_k^{1/4}e^{-2n_k^{1/4}} - \sum_{t=n}^{\infty} t^{1/4}e^{-2t^{1/4}} - 1/r^{11/10} - 1/r^{6/5}.$$

Since

$$\sum_{r=1}^{\infty} \left( 4r^{-91/90} + (\log n_0)n_0^{1/2}e^{-2n_0^{1/25}} + 10^5 n_k^{1/4}e^{-2n_k^{1/4}} + \sum_{t=n}^{\infty} t^{1/4}e^{-2t^{1/4}} + r^{-11/10} + r^{-6/5} \right)$$

is finite, the first Borel-Cantelli Lemma gives us the required result. $\qquad\square$

Notice that in this proof we use Markov's inequality twice, each time introducing a factor of $r$, which is why our bound is (essentially) $|I_r| \leq r^{5+\varepsilon}$ and not $|I_r| \leq r^{3+\varepsilon}$ as we believe.

Note that Theorem 2.14 implies that, almost surely, $|I_r|$ is finite for all $r$, as follows. Suppose for a contradiction that the event that there exists some $x$ with $|I_x|$ infinite has positive probability. Since the probability that $r$ selects $x$ is equal to $2/r$ for $r > x$, we have that $x$ is selected infinitely often, almost surely. So there are an infinite number of elements comparable to $x$ and any such element $r$ must be incomparable with the elements in $I_x \setminus [r]$, meaning that $|I_r| \geq |I_x \setminus [r]|$. Therefore, conditioned on $x$ having $|I_x|$ infinite, we have an infinite number of elements $r$ with $|I_r|$ infinite, almost surely, which contradicts Theorem 2.14.

$\cdot$

## 2.3  A poset not contained in $B_2$

In Section 2.1 we have shown that $B_2$ contains $P(1, 2; m)$ almost surely. It is natural to ask whether this is typical: which posets are contained in $B_2$? For any poset $P$,

{1,2}   {1,3}   {2,3}

{1}     {2}     {3}

Figure 2.1:  $P(1,2;3)$

$k$
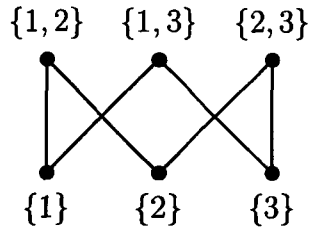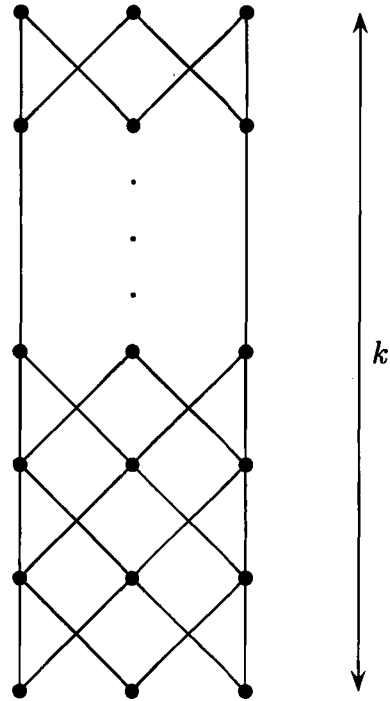
Figure 2.2:  $P(1,2;3)^{(k)}$

$\mathbb{P}(B_2 \supseteq P)$ is positive, as $P$ is a subposet of some possible binary order. So, is every finite poset contained, almost surely? This has been shown for random graph orders; here we show that it is not true for $B_2$.

Recall that we write $P(1,2;3)$ for the poset consisting of the 1-element and 2-element subsets of $\{1,2,3\}$ ordered by inclusion (Figure 2.1). Write $P(1,2;3)^{(k)}$ for a "tower" of $k$ copies of $P(1,2;3)$ with the maximal elements of copy $i$ identified with the minimal elements of copy $i+1$, for $i = 1,\ldots,k-1$ (Figure 2.2).

The result from Proposition 2.2, for the case $m = 3$, is that a copy of $P(1,2;3)$ with minimal points $r, r+1, r+2$ is contained in $B_2[r,n]$, where $n = 2r^{7/5}$, with probability at least $3/5$. The method used certainly requires $k^2 = |U_r|^2 > n = 2r\sqrt{k} + 2k$, i.e., $n \gtrsim r^{4/3}$. We now consider the probability that there exists any copy of $P(1,2;3)^{(k)}$ in $B_2[r,n]$, and show this is very small for $n = o(r^{(k+2)/3})$. (So for $k = 1$ this is a trivial result but, interestingly, if we restrict to only copies of $P(1,2;3)^{(k)}$ with minimal points $r, r+1, r+2$ then the result becomes that the

probability that there exists such a copy in $B_2[r, n]$ is very small for $n = o(r^{k/3+1})$. This gives a certain justification to the method used to construct such a $P(1, 2; 3)$.) Using this result with Theorem 2.14 we provide an example of a poset that, with positive probability, is not contained in $B_2$.

**Theorem 2.15.** *The probability that there exists a $P(1, 2; 3)^{(k)}$ as a subposet of $B_2[r, n]$ is $O(n^9/r^{3k+6})$.*

**Proof.** The proof strategy is as follows. We first define a framework which is a subset of $B_2[r, n]$ satisfying certain properties. The definition of a framework implies that if $B_2[r, n]$ contains no frameworks then it contains no copies of $P(1, 2; 3)^{(k)}$. We then calculate the expected number of frameworks in $B_2[r, n]$ by a path counting method similar to that in the proof of Proposition 2.2. This method provides an upper bound on the expected number of frameworks. The bulk of the proof is in defining a framework in a way that makes the path counting possible. We start with some observations of the structure of copies of $P(1, 2; 3)$ and $P(1, 2; 3)^{(k)}$ in $B_2$, motivating the precise definition of a framework.

Throughout we will write $x$ is above (below) $y$ to mean $x$ is above (below) $y$ in $B_2$, and write $x$ is greater (less) than $y$ to mean $x$ is greater (less) than $y$ in $\mathbb{N}$. Usually, we will reserve $<, \leq, >$ and $\geq$ for the order on $\mathbb{N}$.

Consider $P(1, 2; 3)$ as a subposet of $B_2$ and take a minimal point, $a$. It is below two maximal points, $b_1, b_2$, so there is at least one path from $a$ to $b_1$ and at least one path from $a$ to $b_2$. Choosing one path to $b_1$ and one to $b_2$, we can find the greatest point common to both paths, call this a *branching point*. We can do this for all three minimal points to obtain three branching points. The six chosen paths can also be paired according to which maximal point they go to, and taking the least point common to a pair of paths gives three *connection points*, one for each maximal point. Note that the branching and connection points are not unique if we had a choice of paths, but are distinct for any choice of paths. We label the

(a) $\gamma < \alpha'$        (b) $\alpha' < \gamma$
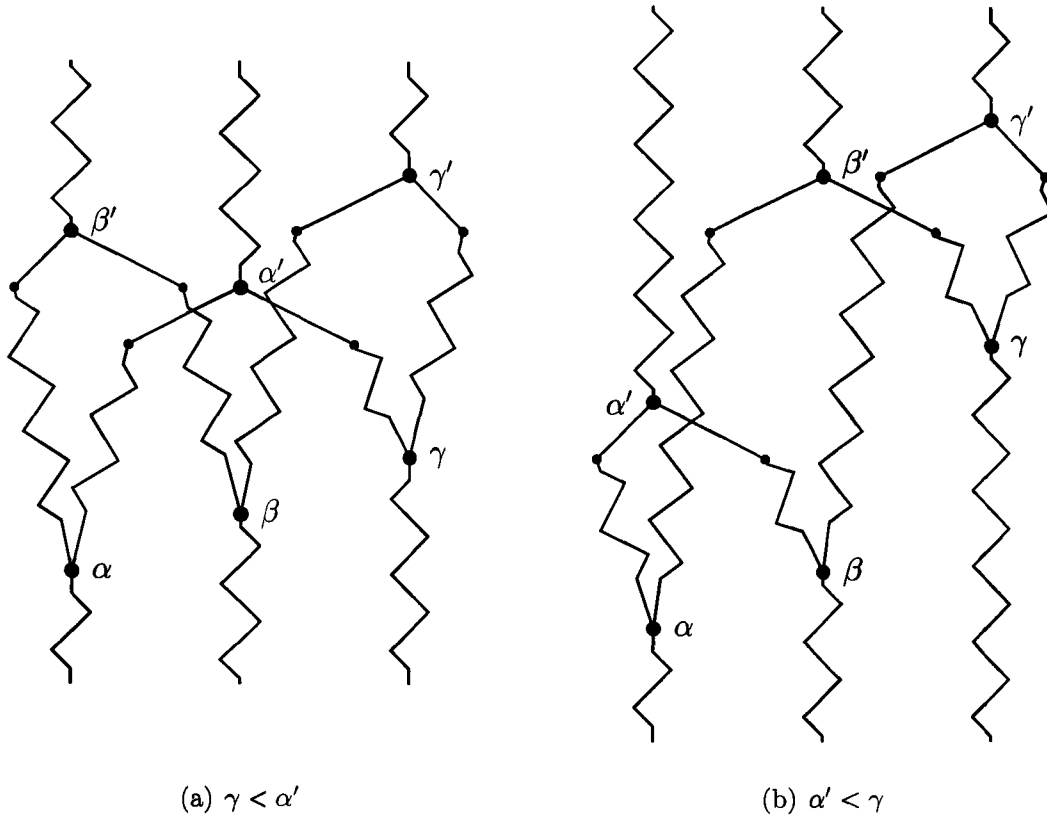
Figure 2.3: $P(1, 2; 3)$ with branching and connection points

branching points $\alpha, \beta, \gamma$ and the connection points $\alpha', \beta', \gamma'$, so that $\alpha < \beta < \gamma$ and $\alpha' < \beta' < \gamma'$. Each path contains both a branching point and a connection point, and since each connection point is contained in two paths, it must be greater than (at least) two branching points. In particular, $\alpha'$ must be greater than $\alpha$ and $\beta$. Similarly, each branching point is less than (at least) two connection points, so $\gamma$ must be less than $\beta'$ and $\gamma'$. So, we have the inequalities $\beta < \alpha'$ and $\gamma < \beta'$, which gives the order $\alpha < \beta < \gamma, \alpha' < \beta' < \gamma'$. It is not possible to order $\gamma$ and $\alpha'$. An example of the branching and connection points for the two cases $\gamma < \alpha'$ and $\alpha' < \gamma$ are shown in Figure 2.3. Note that in Fig. 2.3(a) $\alpha'$ can be above any pair of branching points, whereas in Fig. 2.3(b) $\alpha'$ has to be above $\alpha$ and $\beta$.

For a particular copy of $P(1, 2; 3)^{(k)}$ in $B_2$ we have $k$ copies of $P(1, 2; 3)$ so we can find branching points and connection points for each copy. We label the branching points in copy $i$ by $\alpha_i, \beta_i, \gamma_i$ and the connection points by $\alpha'_i, \beta'_i, \gamma'_i$. So, we have

sequences $\alpha, \beta, \gamma$ of branching points and sequences $\alpha', \beta', \gamma'$ of connection points, where subscript $i$ denotes the points in copy $i$. We have the order $\alpha_i < \beta_i < \gamma_i, \alpha'_i < \beta'_i < \gamma'_i$ for each $i$, as before. Call the points $\alpha_i, \beta_i, \gamma_i$, $i$-branching points, and the points $\alpha'_i, \beta'_i, \gamma'_i$, $i$-connection points.

Ideally, we would aim to separate the copies of $P(1, 2; 3)$ to analyse them individually (for example by assuming $\gamma'_i < \alpha_{i+1}$). Unfortunately this is not possible so we have more cases to consider.

Since $P(1, 2; 3)^{(k)}$ is formed by identifying maximal points in copy $i$ of $P(1, 2; 3)$ to minimal points in copy $i+1$, we have that each $(i+1)$-branching point $\alpha_{i+1} < \beta_{i+1} < \gamma_{i+1}$ is above (and therefore greater than) a distinct $i$-connection point $\alpha'_i < \beta'_i < \gamma'_i$. This immediately gives the inequalities $\alpha_{i+1} > \alpha'_i$ and $\gamma'_i < \gamma_{i+1}$. Looking at $\beta_{i+1}$, either it is above $\beta'_i$ or $\gamma'_i$ which implies $\beta_{i+1} > \beta'_i$, or it is above $\alpha'_i$ in which case $\alpha_{i+1}$ is not above $\alpha'_i$ and so must be above $\beta'_i$ or $\gamma'_i$. But this implies $\beta_{i+1} > \alpha_{i+1} > \beta'_i$. To summarise, we have

$$\alpha_i < \beta_i < \gamma_i, \alpha'_i < \beta'_i < \gamma'_i \qquad \text{for } i = 1, \ldots, k \tag{2.28}$$

$$\alpha_{i+1} > \alpha'_i \qquad \beta_{i+1} > \beta'_i \qquad \gamma_{i+1} > \gamma'_i \qquad \text{for } i = 1, \ldots, k-1 \tag{2.29}$$

which is all we can deduce about the order of branching and connection points.

Suppose we have a $P(1, 2; 3)^{(k)}$ in $B_2[r, n]$. We partition $[r, n]$ into sets of two types (plus two 'end' sets). A set of Type I is of the form $[\beta_i, \beta'_i]$ and a set of Type II of the form $[\beta'_i + 1, \beta_{i+1} - 1]$. The $k$ sets of Type I and $k-1$ sets of Type II and the 'end' sets $[r, \beta_1 - 1]$ and $[\beta'_k + 1, n]$ form the partition of $[r, n]$. We investigate which parts can contain the branching and connection points. Clearly, $\beta_i$ and $\beta'_i$ are contained in the Type I sets. From (2.28) we have that $\gamma_i, \alpha'_i \in [\beta_i, \beta'_i]$ $(i = 1, \ldots, k)$. Also, (2.28) and (2.29) give the inequalities $\beta_{i-1} < \alpha_i < \beta_i$ and $\beta'_i < \gamma'_i < \beta'_{i+1}$ which implies that $\alpha_i \in [\beta_{i-1}, \beta'_{i-1}] \cup [\beta'_{i-1} + 1, \beta_i - 1]$ $(i = 2, \ldots, k)$ and $\gamma'_i \in [\beta'_i + 1, \beta_{i+1} - 1] \cup [\beta_{i+1}, \beta'_{i+1}]$ $(i = 1, \ldots, k-1)$. The end cases $\alpha_1 \in [r, \beta_1 - 1]$ and $\gamma'_k \in [\beta'_k + 1, n]$ are obvious. So, looking at a Type I set $[\beta_i, \beta'_i]$, it contains $\beta_i, \gamma_i, \alpha'_i$ and $\beta'_i$ and possibly $\gamma'_{i-1}$ and

(a) Contains $\beta_i, \gamma_i, \alpha_i', \beta_i'$

(b) Also contains $\gamma_{i-1}'$

(c) Also contains $\alpha_{i+1}$

(d) Also contains $\gamma_{i-1}'$ and $\alpha_{i+1}$
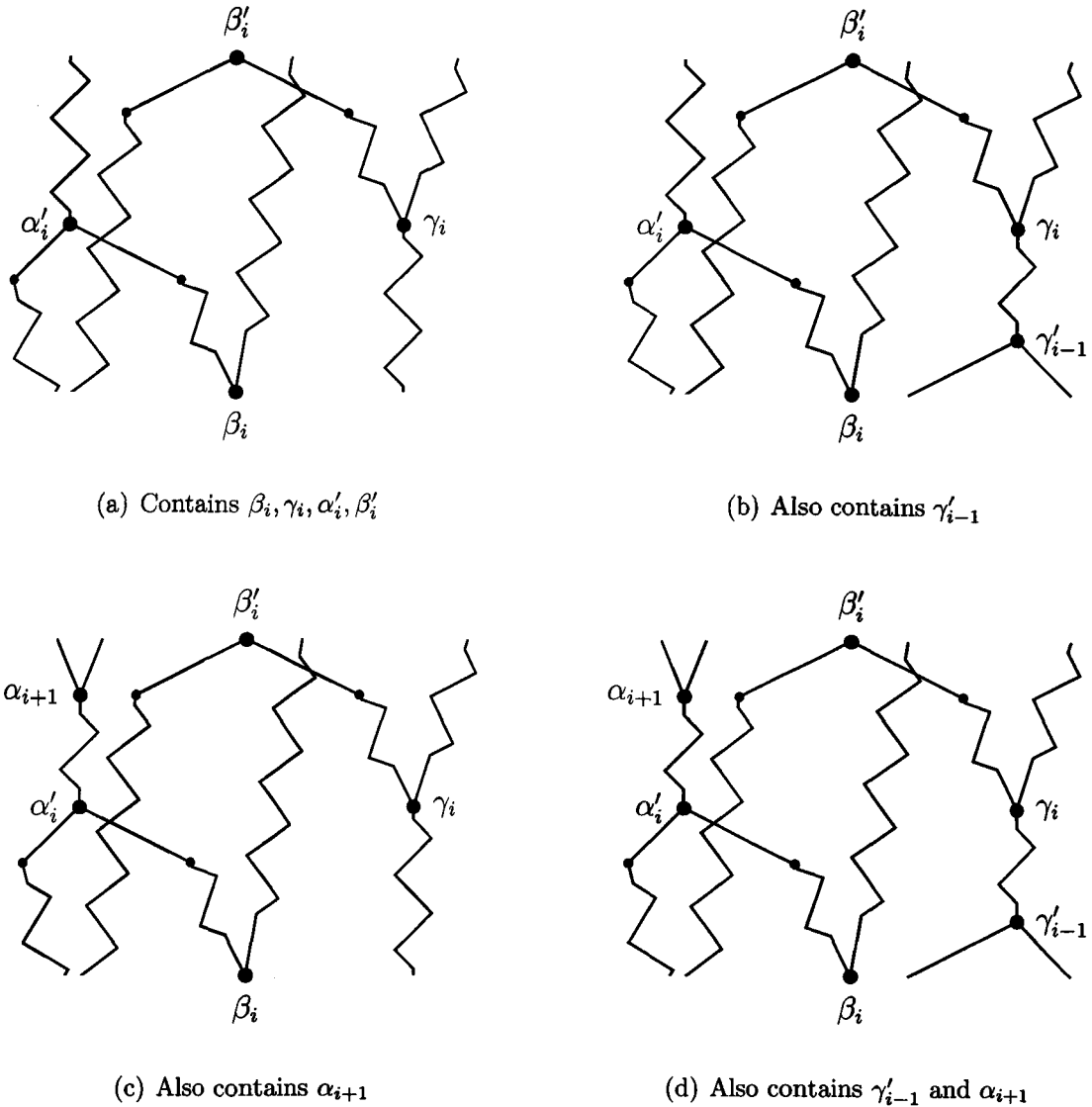
Figure 2.4: Points in $[\beta_i, \beta_i']$ — 4 possible cases

$\alpha_{i+1}$. This gives four possibilities which are shown in Figure 2.4. Finally, we have that the points in the Type II sets are determined by the points in the two adjacent Type I sets. That is, $[\beta_i' + 1, \beta_{i+1} - 1]$ may contain $\gamma_i'$ (but only if $\gamma_i' \notin [\beta_{i+1}, \beta_{i+1}']$) and $\alpha_{i+1}$ (but only if $\alpha_{i+1} \notin [\beta_i, \beta_i']$).

Fix $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$. We call a set $J \subseteq [r, n]$ an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework in $B_2[r, n]$ if $J$ contains all the points in the sequences $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ and the remaining points in $J$ form disjoint paths so that:

(a) there are two paths from each branching point,

(b) there are two paths to each $i$-connection point, which are from two $i$-branching points so that no two $i$-connection points have their paths from the same two $i$-branching points, for $i = 1, \ldots, k$,

(c) there is one path from each connection point (except for the $k$-connection points),

(d) there is one path to each $i$-branching point, which is from a $(i-1)$-connection point, for $i = 2, \ldots, k$.

Note that these paths can just consist of start and end points, that is, it is possible for the set that only contains the points in $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ to be an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework. Indeed, for any set $J \subseteq [r, n]$ containing all the points in $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ there is a positive probability of $J$ being an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework.

For any copy of $P(1, 2; 3)^{(k)}$ with branching points given by $\alpha, \beta, \gamma$ and connection points given by $\alpha', \beta', \gamma'$, we can construct an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework by taking the set of all the branching and connection points and the points of the paths that defined them (but not including those paths below 1-branching points, and those paths above the $k$-connection points). Calling a set $J \subseteq [r, n]$ a *framework* in $B_2[r, n]$ if it is an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework in $B_2[r, n]$ for some $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$, we have that if $B_2[r, n]$ contains no frameworks then it also contains no copies of $P(1, 2; 3)^{(k)}$.

So, it is enough to show that the expected number of frameworks in $B_2[r, n]$ is small and we do this by showing that the expected number of $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-frameworks in $B_2[r, n]$ is small for all sequences $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ satisfying (2.28) and (2.29).

For fixed $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$, we count the number of $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-frameworks in $B_2[r, n]$ by considering the event "$J$ is an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework" as a sequence of events in the sets of the partition of $[r, n]$. That is, we split an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework into $j$-frameworks and $l$-frames, defined below, and

we show that it is possible to count the expected number of $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-frameworks by independently counting the number of $l$-frames in each part of the partition.

Label the partition

$$K_1 = [r, \beta_1 - 1], \quad K_{2k+1} = [\beta'_k + 1, n]$$

$$K_{2i} = [\beta_i, \beta'_i], i = 1, \ldots, k$$

$$K_{2i+1} = [\beta'_i + 1, \beta_{i+1} - 1], i = 1, \ldots, k - 1.$$

We write $\max K_j$ for the largest element of $K_j$. In a definition similar to that of an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework, for $j = 1, \ldots, 2k + 1$, we call a set $J \subseteq [r, \max K_j]$ a $j$-framework in $B_2[r, \max K_j]$ if $J$ contains all the points in the sequences $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ that are in $[r, \max K_j]$ and the remaining points in $J$ form disjoint paths so that

(a) there are two paths from each branching point in $J$,

(b) there are two paths to each $i$-connection point in $J$, which are from two $i$-branching points in $J$ so that no two $i$-connection points in $J$ have their paths from the same two $i$-branching points in $J$, for $i = 1, \ldots, k$,

(c) there is one path from each connection point in $J$ (except for the $k$-connection points),

(d) there is one path to each $i$-branching point in $J$, which is from a $(i-1)$-connection point in $J$, for $i = 2, \ldots, k$.

Again, for any set $J \subseteq [r, \max K_j]$ containing all the points in $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ that are in $[r, \max K_j]$ there is a positive probability of $J$ being a $j$-framework.

So, a $(2k+1)$-framework is the same as an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework. Notice that, whereas in a $(2k+1)$-framework all paths are between branching and connection

points, in a $j$-framework, for $j \neq 2k + 1$, there can be paths from some branching and connection points that do not end at a branching or connection point (the paths from the branching and connection points that are not below any others in $J$). Call the end points of these paths the *end points* of the $j$-framework. We shall see that, although the end points of a $j$-framework can be different for different $j$-frameworks, what is important for our calculations is that the number of end points of a $j$-framework is the same for different $j$-frameworks, for fixed $j$.

Now, define an $l$-frame as follows:

- $l = 1$: A 1-*frame* is a set $J_1 \subseteq K_1$ which is a 1-framework in $B_2[r, \max K_1]$.

- $l \neq 1$: Given that $J$ is an $(l-1)$-framework in $B_2[r, \max K_{l-1}]$, an $l$-*frame for* $J$ is a set $J_l \subseteq K_l$ such that $J \cup J_l$ is an $l$-framework in $B_2[r, \max K_l]$.

So, for sets $J_j \subseteq K_j$, $j = 1, \ldots, 2k + 1$, we have

$$
\mathbb{P} \left( \bigcup_{j=1}^{2k+1} J_j \text{ an } (\alpha, \beta, \gamma, \alpha', \beta', \gamma')\text{-framework} \right) =
$$

$$
\mathbb{P}(J_1 \text{ a 1-frame}) \mathbb{P}(J_2 \text{ a 2-frame for } J_1) \cdots \mathbb{P} \left( J_{2k+1} \text{ a } (2k+1)\text{-frame for } \bigcup_{j=1}^{2k} J_j \right).
$$

$$(2.30)$$

Now, write $X(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$ for the number of $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-frameworks. We have $X(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$ equal to the sum

$$
\sum_{J_1 \subseteq K_1} \cdots \sum_{J_{2k+1} \subseteq K_{2k+1}} I \left( \bigcup_{j=1}^{2k+1} J_j \text{ is an } (\alpha, \beta, \gamma, \alpha', \beta', \gamma')\text{-framework} \right), \qquad (2.31)
$$

but $\bigcup_{j=1}^{2k+1} J_j$ is an $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$-framework only if $\bigcup_{j=1}^{2k+1} J_j$ contains all the points in $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$. So, writing $K_j(BC)$ for the set of branching and connection points that are in $K_j$, the sum (2.31) is equal to

$$
\sum_{\substack{J_1 \subseteq K_1: \\ K_1(BC) \subseteq J_1}} \cdots \sum_{\substack{J_{2k+1} \subseteq K_{2k+1}: \\ K_{2k+1}(BC) \subseteq J_{2k+1}}} I \left( \bigcup_{j=1}^{2k+1} J_j \text{ is an } (\alpha, \beta, \gamma, \alpha', \beta', \gamma')\text{-framework} \right).
$$

Taking expectations and using (2.30) gives

$$\mathbb{E}X(\alpha, \beta, \gamma, \alpha', \beta', \gamma') =$$

$$\sum_{\substack{J_1 \subseteq K_1: \\ K_1(BC) \subseteq J_1}} \cdots \sum_{\substack{J_{2k+1} \subseteq K_{2k+1}: \\ K_{2k+1}(BC) \subseteq J_{2k+1}}} \mathbb{P}(J_1 \text{ a 1-frame}) \cdots \mathbb{P}\left(J_{2k+1} \text{ a } (2k+1)\text{-frame for } \bigcup_{j=1}^{2k} J_j\right)$$

$$(2.32)$$

But $\mathbb{P}\left(J_l \text{ an } l\text{-frame for } \bigcup_{j=1}^{l-1} J_j\right)$ does not depend on $J_1, \ldots, J_{l-1}$; this is the conditional probability that $\bigcup_{j=1}^{l} J_j$ is an $l$-framework, given that $\bigcup_{j=1}^{l-1} J_j$ is an $(l-1)$-framework. Since $K_l(BC) \subseteq J_l$, this is the probability that the points in $\bigcup_{j=1}^{l} J_j$ form paths satisfying (a)–(d). But we know that $\bigcup_{j=1}^{l-1} J_j$ is an $(l-1)$-framework, so $\bigcup_{j=1}^{l} J_j$ is an $l$-framework provided the points in $J_l$ form paths that continue the paths in $\bigcup_{j=1}^{l-1} J_j$ in such a way that (a)–(d) are satisfied. That is, the points in $J_l$ must either select other points in $J_l$, or one of the end points of the $(l-1)$-framework, $\bigcup_{j=1}^{l-1} J_j$. So the probability $\mathbb{P}\left(J_l \text{ an } l\text{-frame for } \bigcup_{j=1}^{l-1} J_j\right)$ can only depend on the set $J_l$ and the number of end points of $\bigcup_{j=1}^{l-1} J_j$. However, the number of end points of a $j$-framework is determined by which branching and connection points are not below any others in the $j$-framework and these are fixed for particular sequences $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$.

So, for $j = 2, \ldots, 2k+1$ we write $\mathbb{P}_l(J_l)$ for $\mathbb{P}\left(J_l \text{ an } l\text{-frame for } \bigcup_{j=1}^{l-1} J_j\right)$, and we write $\mathbb{P}_1(J_1)$ for $\mathbb{P}\left(J_1 \text{ a 1-frame}\right)$. Equation (2.32) becomes

$$\mathbb{E}X(\alpha, \beta, \gamma, \alpha', \beta', \gamma') = \prod_{l=1}^{2k+1} \sum_{\substack{J_l \subseteq K_l: \\ K_l(BC) \subseteq J_l}} \mathbb{P}_l(J_l)$$

Writing $X$ for the total number of frameworks and $\mathbb{E}_l$ for $\sum_{J_l \subseteq K_l : K_l(BC) \subseteq J_l} \mathbb{P}_l(J_l)$, we have that the expected number of frameworks is

$$\mathbb{E}X = \sum_{\substack{\alpha, \beta, \gamma, \\ \alpha', \beta', \gamma' :(2.28),(2.29)}} \prod_{l=1}^{2k+1} \mathbb{E}_l.$$

We now calculate an upper bound for each $\mathbb{E}_l$ by a path counting method. There are various cases to consider depending on the ordering of the branching and connec-

tion points $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$. However, we calculate an upper bound for $\prod_{l=1}^{2k+1} \mathbb{E}_l$ for the case $\alpha_i < \beta_i < \gamma_i < \alpha_i' < \beta_i' < \gamma_i' < \alpha_{i+1}$, etc. (Figure 2.4(a)) and then show that this ordering is the worst case. That is, that the upper bound for the case $\alpha_i < \beta_i < \gamma_i < \alpha_i' < \beta_i' < \gamma_i' < \alpha_{i+1}$, etc, is an upper bound for any ordering of the branching and connection points subject to (2.28) and (2.29).

We again use the inequalities $1 + x \leq e^x$ and $\sum_{j=a}^{b} f(j) \leq \int_{a-1}^{b} f(x)dx$ for $f(x)$ decreasing, so that in particular

$$\prod_{j=a}^{b} \left(1 + \frac{c}{j}\right) \leq \exp\left(\sum_{j=a}^{b} \frac{c}{j}\right) \leq \exp\left(c \log \frac{b}{a-1}\right) = \left(\frac{b}{a-1}\right)^c.$$

For $l = 1$, $K_1(BC) = \{\alpha_1\}$, so we sum over $J_1 \subseteq K_1 = [r, \beta_1 - 1]$ containing $\{\alpha_1\}$. If $J_1 = \{\alpha_1, j_1, \ldots, j_t\}$ with $\alpha_1 < j_1 < \cdots < j_t \leq \beta_1 - 1$, then the probability $\mathbb{P}_1(J_1)$ is the probability that the points $j_s, s = 1, \ldots, t$ form two disjoint paths from $\alpha_1$, which is at most $\prod_{s=1}^{t}(4/j_s)$, by independence, and if $J_1 \not\subseteq [\alpha_1, \beta_1 - 1]$ then $\mathbb{P}_1(J_1) = 0$, so

$$\mathbb{E}_1 \leq \prod_{j=\alpha_1+1}^{\beta_1-1} \left(1 + \frac{4}{j}\right) \leq \left(\frac{\beta_1 - 1}{\alpha_1}\right)^4 \leq \frac{\beta_1^4}{\alpha_1^4}.$$

For $l = 2$, $K_2 = \{\beta_1, \gamma_1, \alpha_1', \beta_1'\}$, we sum over $J_2 \subseteq K_2 = [\beta_1, \beta_1']$ containing $\{\beta_1, \gamma_1, \alpha_1', \beta_1'\}$. So, $J_2 = \{\beta_1, j_1^{(1)}, \ldots, j_{t_1}^{(1)}, \gamma_1, j_1^{(2)}, \ldots, j_{t_2}^{(2)}, \alpha_1', j_1^{(3)}, \ldots, j_{t_3}^{(3)}, \beta_1'\}$ and the probability $\mathbb{P}_2(J_2)$ is the probability that

(i) the points $j_s^{(1)}, s = 1, \ldots, t_1$ form four disjoint paths — two from $\beta_1$, two from the existing end points in the 1-frame $J_1$,

(ii) the points $j_s^{(2)}, s = 1, \ldots, t_2$ form six disjoint paths — two from $\gamma_1$, four from the end points of the paths formed in (i),

(iii) the point $\alpha_1'$ is above two of the end points of the paths formed in (ii) (specifically, two paths with different starting points),

(iv) the points $j_s^{(3)}, s = 1, \ldots, t_3$ form five disjoint paths — one from $\alpha_1'$, four from the end points of the remaining paths formed in (ii),

(v) the point $\beta_1'$ is above two of the end points of the four "branching" paths formed in (iv) (i.e., not the path from $\alpha_1'$, and again specifically, two paths with different starting points).

All these events are independent of each other, and so this probability is at most

$$\left(\prod_{s=1}^{t_1} \frac{8}{j_s^{(1)}}\right)\left(\prod_{s=1}^{t_2} \frac{12}{j_s^{(2)}}\right)\frac{12}{\binom{\alpha_1'}{2}}\left(\prod_{s=1}^{t_3} \frac{10}{j_s^{(3)}}\right)\frac{4}{\binom{\beta_1'}{2}},$$

so the sum over all subsets of $K_2$ is

$$\mathbb{E}_2 \leq \prod_{j=\beta_1+1}^{\gamma_1-1}\left(1+\frac{8}{j}\right)\prod_{j=\gamma_1+1}^{\alpha_1'-1}\left(1+\frac{12}{j}\right)\frac{12}{\binom{\alpha_1'}{2}}\prod_{j=\alpha_1'+1}^{\beta_1'-1}\left(1+\frac{10}{j}\right)\frac{4}{\binom{\beta_1'}{2}}$$

$$\leq \left(\frac{\gamma_1-1}{\beta_1}\right)^8\left(\frac{\alpha_1'-1}{\gamma_1}\right)^{12}\frac{24}{\alpha_1'(\alpha_1'-1)}\left(\frac{\beta_1'-1}{\alpha_1'}\right)^{10}\frac{8}{\beta_1'(\beta_1'-1)}$$

$$\leq 2^6 3\frac{\beta_1'^8}{\beta_1^8\gamma_1^4}.$$

For $l = 2i+1$ $(i = 1,\ldots,k-1)$, $K_{2i+1} = [\beta_i'+1, \beta_{i+1}-1]$, $K_{2i+1}(BC) = \{\gamma_i', \alpha_{i+1}\}$ and by a similar calculation we have

$$\mathbb{E}_{2i+1} \leq \prod_{j=\beta_i'+1}^{\gamma_i'-1}\left(1+\frac{8}{j}\right)\frac{1}{\binom{\gamma_i'}{2}}\prod_{j=\gamma_i'+1}^{\alpha_{i+1}-1}\left(1+\frac{6}{j}\right)\frac{6}{\alpha_{i+1}}\prod_{j=\alpha_{i+1}+1}^{\beta_{i+1}-1}\left(1+\frac{8}{j}\right)$$

$$\leq \left(\frac{\gamma_i'-1}{\beta_i'}\right)^8\frac{2}{\gamma_i'(\gamma_i'-1)}\left(\frac{\alpha_{i+1}-1}{\gamma_i'}\right)^6\frac{6}{\alpha_{i+1}}\left(\frac{\beta_{i+1}-1}{\alpha_{i+1}}\right)^8$$

$$\leq 2^2 3\frac{\beta_{i+1}^8}{\beta_i'^8\alpha_{i+1}^3}$$

and for $l = 2i$ $(i = 2,\ldots,k-1)$, $K_{2i} = [\beta_i, \beta_i']$, $K_{2i}(BC) = \{\beta_i, \gamma_i, \alpha_i', \beta_i'\}$ and

$$\mathbb{E}_{2i} \leq \frac{4}{\beta_i}\prod_{j=\beta_i+1}^{\gamma_i-1}\left(1+\frac{10}{j}\right)\frac{2}{\gamma_i}\prod_{j=\gamma_i+1}^{\alpha_i'-1}\left(1+\frac{12}{j}\right)\frac{12}{\binom{\alpha_i'}{2}}\prod_{j=\alpha_i'+1}^{\beta_i'-1}\left(1+\frac{10}{j}\right)\frac{4}{\binom{\beta_i'}{2}}$$

$$\leq \frac{4}{\beta_i}\left(\frac{\gamma_i-1}{\beta_i}\right)^{10}\frac{2}{\gamma_i}\left(\frac{\alpha_i'-1}{\gamma_i}\right)^{12}\frac{24}{\alpha_i'(\alpha_i'-1)}\left(\frac{\beta_i'-1}{\alpha_i'}\right)^{10}\frac{8}{\beta_i'(\beta_i'-1)}$$

$$\leq 2^9 3\frac{\beta_i'^8}{\beta_i^{11}\gamma_i^3},$$

and

$$\mathbb{E}_{2k} \leq \frac{4}{\beta_k}\prod_{j=\beta_k+1}^{\gamma_k-1}\left(1+\frac{10}{j}\right)\frac{2}{\gamma_k}\prod_{j=\gamma_k+1}^{\alpha_k'-1}\left(1+\frac{12}{j}\right)\frac{12}{\binom{\alpha_k'}{2}}\prod_{j=\alpha_k'+1}^{\beta_k'-1}\left(1+\frac{8}{j}\right)\frac{4}{\binom{\beta_k'}{2}}$$

$$\leq \frac{4}{\beta_k}\left(\frac{\gamma_k-1}{\beta_k}\right)^{10}\frac{2}{\gamma_k}\left(\frac{\alpha_k'-1}{\gamma_k}\right)^{12}\frac{24}{\alpha_k'(\alpha_k'-1)}\left(\frac{\beta_k'-1}{\alpha_k'}\right)^8\frac{8}{\beta_k'(\beta_k'-1)}$$

$$\leq 2^9 3\frac{\alpha_k'^2\beta_k'^6}{\beta_k^{11}\gamma_k^3}$$

and

$$\mathbb{E}_{2k+1} \leq \prod_{j=\beta_k'+1}^{\gamma_k'-1} \left(1 + \frac{4}{j}\right) \frac{1}{\binom{\gamma_k'}{2}} \leq \left(\frac{\gamma_k'-1}{\beta_k'}\right)^4 \frac{2}{\gamma_k'(\gamma_k'-1)} \leq 2\frac{\gamma_k'^2}{\beta_k'^4}.$$

This gives the upper bound

$$\prod_{l=1}^{2k+1} \mathbb{E}_l \leq \frac{\beta_1^4}{\alpha_1^4} . 2^6 3 \frac{\beta_1'^8}{\beta_1^8 \gamma_1^4} \prod_{i=1}^{k-1} \left(2^2 3 \frac{\beta_{i+1}^8}{\beta_i'^8 \alpha_{i+1}^3}\right) \prod_{i=2}^{k-1} \left(2^9 3 \frac{\beta_i'^8}{\beta_i^{11} \gamma_i^3}\right) 2^9 3 \frac{\alpha_k'^2 \beta_k'^6}{\beta_k^{11} \gamma_k^3} . 2 \frac{\gamma_k'^2}{\beta_k'^4}$$

$$= 2^7 3 \frac{\alpha_k'^2 \beta_k'^2 \gamma_k'^2}{\alpha_1^4 \beta_1^4 \gamma_1^4} \prod_{i=2}^{k} \frac{2^{11} 3^2}{\alpha_i^3 \beta_i^3 \gamma_i^3}.$$

We show that this is also an upper bound on $\mathbb{E}X(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$ for any ordering of the branching and connection points $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$. For any other ordering, where some of the $\alpha_{i+1}$ and $\gamma_{i-1}'$ fall into $K_{2i}$, we can carry out a similar calculation, and obtain an expression of a similar form, namely

$$A_k \prod_{i=1}^{k} (\alpha_i \beta_i \gamma_i)^{b_i} (\alpha_i' \beta_i' \gamma_i')^{c_i}.$$

For any framework, from the conditions (a)–(d) in the definition, every $i$-branching point ($i \neq 1$) must have one fewer path to it than from it (two fewer for $i = 1$), but $b_i$ depends only on this difference, so $b_i$ is independent of the ordering of the terms of $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$. Similarly, $c_i$ is independent of the ordering since, for any framework, each $i$-connection point ($i \neq k$) has one more path to it than from it (two more for $i = k$). So we have

$$b_i = \begin{cases} 2 \times (-1) - 1 = -3 & \text{for } i \neq 1, \\ 2 \times (-2) \phantom{-1} = -4 & \text{for } i = 1, \end{cases} \qquad c_i = \begin{cases} 2 \times (+1) - 2 = 0 & \text{for } i \neq k, \\ 2 \times (+2) - 2 = 2 & \text{for } i = k, \end{cases}$$

for any ordering. The constant factor, $A_k$, does depend on the ordering of the terms of $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$. In particular, it depends on the number of choices of end points (or pairs of end points) of paths that each branching (or connection) point can be above, respectively. It remains to show that this number is smaller for any ordering (satisfying (2.28) and (2.29)) other than $\alpha_i < \beta_i < \gamma_i < \alpha_i' < \beta_i' < \gamma_i' < \alpha_{i+1}$, etc.

Suppose we have an ordering where $\alpha_i' < \gamma_i$ for some $i$. Then there is only a choice of four pairs of end points of paths for $\alpha_i'$ to be above, rather than the

twelve pairs of end points in the case $\gamma_i < \alpha_i'$. So we need only consider orderings with $\gamma_i < \alpha_i'$ for all $i$. This means events occurring below $\gamma_i$ are independent of events occurring above $\alpha_i'$. In particular we can consider the cases illustrated in Figures 2.4(b) and 2.4(c) separately (so the case in Figure 2.4(d) is a combination of the two). If we have an ordering with $\gamma_{i-1}' > \beta_i$, then there is only one end point for $\beta_i$ to be above, rather than the two end points in the case $\gamma_{i-1}' < \beta_i$. If we have an ordering with $\alpha_{i+1} < \beta_i'$ then there is only one end point $\alpha_{i+1}$ can be above, rather than the two end points in the case $\alpha_{i+1} > \beta_i'$. This only leaves the case that $\alpha_i < \gamma_{i-1}'$, but then there is only a choice of two end points for $\alpha_i$ to be above, rather than the three end points in the case $\alpha_i > \gamma_{i-1}'$.

Therefore,

$$\mathbb{E}X \leq \sum_{\substack{\alpha,\beta,\gamma, \\ \alpha',\beta',\gamma':(2.28),(2.29)}} 2^7 3 \frac{\alpha_k'^2 \beta_k'^2 \gamma_k'^2}{\alpha_1^4 \beta_1^4 \gamma_1^4} \prod_{i=2}^{k} \frac{2^{11}3^2}{\alpha_i^3 \beta_i^3 \gamma_i^3},$$

and summing first over $\alpha_i' < \alpha_{i+1}$ for $i = 1, \ldots, k-1$ (and similarly for $\beta', \gamma'$) and then relaxing all other constraints gives

$$\mathbb{E}X \leq \sum_{\substack{\alpha,\beta,\gamma, \\ \alpha_k',\beta_k',\gamma_k'}} 2^7 3 \frac{\alpha_k'^2 \beta_k'^2 \gamma_k'^2}{\alpha_1^4 \beta_1^4 \gamma_1^4} \prod_{i=2}^{k} \frac{2^{11}3^2}{\alpha_i^2 \beta_i^2 \gamma_i^2}$$

$$\leq \frac{2^7}{3^5} \frac{n^9}{r^9} \left(\frac{2^{11}3^2}{r^3}\right)^{k-1}$$

$$= (2^7/3^5)(2^{11}3^2)^{k-1} \frac{n^9}{r^{3k+6}}.$$

So, the probability that there exists a copy of $P(1,2;3)^{(k)}$ in $B_2[r,n]$ is less than the probability that there exists a framework in $B_2[r,n]$, which is $O(n^9/r^{3k+6})$ by Markov's inequality. $\qquad \square$

We define the poset $Q(k)$ as the poset $P(1,2;3)^{(k)}$ with an additional point incomparable to all others. Write $B_2[r,\infty)$ for the random poset $B_2$ restricted to the set of points greater than or equal to $r$. We have the following corollary of Theorems 2.14 and 2.15.

**Corollary 2.16.** *For $k \geq 450$, the probability that $B_2[r,\infty)$ contains a copy of*

$Q(k-1)$ *is* $O(r^{-91/90})$.

**Proof.** For there to be a copy of $Q(k-1)$ in $B_2[r,\infty)$ there must exist a copy $P$ of $P(1,2;3)^{(k-1)}$ in $B_2[r,\infty)$, and some point $b$ in $B_2[r,\infty)$ such that $b$ is incomparable to all the points in $P$. Label the least point in $P$ by $m$, and the greatest point by $n$, so that $P$ is in $B_2[m,n]$, and $b$ must be incomparable to $m$ and $n$. So, the probability that there is a copy of $Q(k-1)$ in $B_2[r,\infty)$ is less than the probability that there both exists some $P$ in $B_2[m,n]$, and some $b \geq r$ incomparable to $m$ and $n$, for some $m,n \geq r$. If $n = \omega(m^{150})$ then the probability that there exists some $b$ incomparable to both $m$ and $n$ is $O(r^{-91/90})$. Now taking $k \geq 450$, if $n = O(m^{150})$ then the probability there exists an $P$ in $B_2[m,n]$ is $O(m^{-3}) = O(r^{-3})$, since $m \geq r$. So for fixed $k \geq 450$ the probability that $B_2[r,\infty)$ contains a copy of $Q(k-1)$ is $O(r^{-91/90})$. $\square$

Since events in $B_2[r]$ are independent of events in $B_2[r,\infty)$ we have the following corollary.

**Corollary 2.17.** *For $k \geq 450$, there is a positive probability that the random poset $B_2$ does not contain a copy of $Q(k)$.*

**Proof.** Fix $k \geq 450$. Fix $r$ so that the probability that $B_2[r,\infty)$ does not contain a copy of $Q(k-1)$ is at least $1/2$. This is possible by Corollary 2.16.

With some positive probability $p$, the points $2,\ldots,r$ in $B_2$ form a chain. (For this to happen, each point $j = 3,\ldots,r$ must select point $j-1$, so $p = \prod_{j=3}^{r}(2/j) = 2^{r-1}/r!$.) Recall that points 0 and 1 are defined to be incomparable, and vertex 2 selects 0 and 1 with probability 1, so all points in $[r]$ are below $r$ in $B_2[r]$.

Now, we can calculate the probability that $B_2$ contains a copy of $Q(k)$ given that the first $r$ elements are as above. Suppose such a $B_2$ contains a copy $Q$ of $Q(k)$. Because of the structure of $B_2[r]$ there can be at most one point of $Q$ in $B_2[r]$. Either this is the incomparable element of $Q$, or one of the minimal points of the

tower in $Q$. If the former, label this point $b$, and we have $b \leq r$ and so $b$ is below $r$ in $B_2$. The point $b$ is incomparable with all points in $Q$, which implies that $r$ is also incomparable with all points in $Q$. Since $Q$ is a copy of $Q(k)$, so is $Q \cup \{r\} \setminus \{b\}$, and there is a copy of $Q(k)$ in $B_2[r, \infty)$. If the latter, then $Q$ contains a copy of $Q(k-1)$ with all points greater than $r$, that is, a copy of $Q(k-1)$ in $B_2[r, \infty)$. If none of the points in $Q$ are in $B_2[r]$, then $Q$, a copy of $Q(k)$, is contained in $B_2[r, \infty)$.

So, $B_2$ does not contain a copy of $Q(k)$ if $B_2[r, \infty)$ does not contain a copy of $Q(k-1)$. However, the probability of this is at least $1/2$, and is independent of the events in $B_2[r]$. Therefore the probability that $B_2$ does not contain a copy of $Q(k)$ is at least $p/2 > 0$.                                      $\square$

We have shown that there is a positive probability that $B_2$ does not contain $Q(k)$, that is, that $Q(k)$ is not almost surely contained in $B_2$. So, which posets are almost surely contained in $B_2$? It seems ambitious to ask for a complete answer, but it may be possible to provide both families of posets almost surely contained in $B_2$, and families of posets not almost surely contained in $B_2$. We have already shown that $Q(k)$, $k \geq 450$ (and so, also, any posets containing $Q(k)$) are not almost surely contained in $B_2$. In fact, we can apply the argument used in Corollary 2.16 to any poset in place of $P(1, 2; 3)^{(k-1)}$, if we can show that it is not contained in $B_2[r, r^{150}]$ almost surely. This is one way to provide further examples of posets not almost surely contained in $B_2$.

# Chapter 3

# Continuum limits of classical sequential growth models

This chapter describes work carried out in conjunction with my supervisor, Professor Graham Brightwell, and was worked on in equal proportion by myself and Professor Brightwell.

In [25], Rideout and Sorkin provide evidence for a "continuum limit of transitive percolation". Transitive percolation, a model of random partial orders, is specified by one parameter $p$, and produces partial orders sequentially, as follows. We start with a single element, labelled 0. At stage $n = 1, 2, \ldots$, the element $n$ is added to the partial order and placed above each existing element independently with probability $p$, and incomparable to it with probability $1 - p$. The transitive closure of the added relations gives the partial order $P_{n+1,p}$ at stage $n$. From this definition we can see that the poset $P_{n,p}$ is what is called a *random graph order* in the mathematics literature. As mentioned in Chapter 1, these were introduced by Albert and Frieze [1] and have been studied further by Bollobás and Brightwell [7, 8, 9], Pittel and Tungol [23], and Simon, Crippa and Collenberg [27].

In this chapter, we confirm the observation of Rideout and Sorkin, that certain

sequences of random graph orders do have "continuum limits". We also show that, even in a broader class of models, these continuum limits are essentially the only ones that arise.

We start by defining carefully what it means for a sequence $(\mathcal{P}_n)_{n=1}^{\infty}$ of probability spaces, whose elements are finite partial orders, to have an atomless partially ordered measure space as a continuum limit. Usually, the partial orders in $\mathcal{P}_n$ will have ground sets of size $n$.

We use a definition of a partially ordered measure space similar to that in Bollobás and Brightwell [6].

**Definition 3.1.** A *partially ordered measure space* is a quadruple $(X, \mathcal{F}, \mu, <)$ such that $(X, \mathcal{F}, \mu)$ is a measure space, $(X, <)$ is a partially ordered set, and $U[x] \equiv \{y \in X : y \geq x\} \in \mathcal{F}$, and $D[x] \equiv \{y \in X : y \leq x\} \in \mathcal{F}$ for every $x \in X$.

A partially ordered measure space $(X, \mathcal{F}, \mu, <)$ is *atomless* if $\mu(\{x\}) = 0$ for all $x \in X$.

We now give formal definitions of the sampling from partially ordered measure spaces, and the probability of forming a particular labelled partial order $Q$. (In this context, the elements of $Q$ will be labelled $x_1, \ldots, x_k$.)

**Definition 3.2.** For $P$ a partially ordered measure space with probability measure $\mu$, and $k$ a natural number, define a *random sample* of $k$ elements from $P$ to be a sequence $x_1, \ldots, x_k$ of elements of $P$, obtained by selecting $k$ elements $x_1, \ldots, x_k$ independently from $P$ according to $\mu$, and conditioning on the event that $x_1, \ldots, x_k$ are distinct. A random sample can be thought of as a (random) finite partial order on the fixed ground-set $\{x_1, \ldots, x_k\}$, inheriting the partial order from $P$.

For $Q$ a finite partial order with ground-set labelled as $\{x_1, \ldots, x_k\}$, and $P$ a partially ordered measure space with measure $\mu$, let $\lambda(Q; P)$ be the probability that the partial order inherited from $P$ on a random sample $x_1, \ldots, x_k$ of $k$ elements is equal to $Q$.

Note that, for $P$ an atomless partially ordered measure space, the probability that the same element from $P$ is selected twice is zero, and so conditioning on the elements of a random sample being distinct makes no difference.

When we apply the above definitions to a finite partial order $P = (X, <)$, we always take the probability measure $\mu$ to be uniform on $X$. With this convention, sampling $|Q|$ elements from $P$, conditioned on the elements being different, is equivalent to selecting $|Q|$ elements from $P$ without replacement. Therefore $\lambda(Q; P)$ is the proportion of labelled $|Q|$-element subsets of $P$ that are equal to $Q$. To be precise, for $Q$, $P$ finite labelled partial orders, if we select $|Q|$ elements without replacement from $P$, label them with $x_1, \ldots, x_{|Q|}$ according to the order of selection, and take the induced order from $P$, then $\lambda(Q; P)$ is the probability that this random partial order is equal to $Q$.

Note that for fixed $P$, we have $\lambda(Q; P) = \lambda(Q'; P)$ if the labelled posets $Q$ and $Q'$ are isomorphic.

We are now in a position to define a continuum limit. Here, and in what follows, $P_n$ denotes a random partial order from $\mathcal{P}_n$.

**Definition 3.3.** A *continuum limit* of $(\mathcal{P}_n)_{n=1}^\infty$, a sequence of probability spaces, whose elements are finite posets, is an atomless partially ordered measure space $P_\infty$ such that, for all finite labelled partial orders $Q$,

$$\mathbb{E}\lambda(Q; P_n) \to \lambda(Q; P_\infty).$$

In [25], Rideout and Sorkin estimate $\lambda(Q; P_{n,p})$ for small partial orders $Q$, and present evidence suggesting that, for suitable sequences $p = p(n)$, all the expectations $\mathbb{E}\lambda(Q; P_{n,p})$ converge to limits. To be more precise, they choose sequences $p(n)$ so that $\mathbb{E}\lambda(C_2; P_{n,p})$ converges, where $C_2$ is the 2-element chain, and observe that, for such sequences $p(n)$, expectations $\mathbb{E}\lambda(Q; P_{n,p})$, for other small $Q$, appear to converge also. They offer this as evidence for the existence of a continuum limit.

We define a sequence $(\mathcal{P}_n)_{n=1}^{\infty}$ of discrete probability spaces to be *compatible* if $\left(\mathbb{E}\lambda(Q; P_n)\right)_{n=1}^{\infty}$ is convergent for all finite labelled partial orders $Q$. From the definitions we have that, if $(\mathcal{P}_n)_{n=1}^{\infty}$ has a continuum limit, then $(\mathcal{P}_n)_{n=1}^{\infty}$ is compatible. An interesting question (not answered here) is whether a compatible sequence necessarily has a continuum limit. In Section 3.1.2, we show not only that suitable sequences of random graph orders are compatible but also that they have continuum limits, confirming the conjecture of Rideout and Sorkin.

**Theorem 3.4.** *The sequence of models* $(\mathcal{P}_{n,p(n)})_{n=1}^{\infty}$ *of random graph orders has a continuum limit if and only if one of the following holds:*

*(i)* $\lim_{n\to\infty} \left(p^{-1}\log p^{-1}/n\right) = 0,$

*(ii)* $\lim_{n\to\infty} \left(p^{-1}\log p^{-1}/n\right) = c$ *for some* $0 < c < 1,$ *or*

*(iii)* $\liminf_{n\to\infty} \left(p^{-1}\log p^{-1}/n\right) \geq 1.$

In the first and third of the cases above, the continuum limit is very trivial, being a chain and an antichain respectively. In the second case, the continuum limit is a "random semiorder".

A *semiorder* is a partial order that can be represented by a collection of equal-length intervals on the real line, ordered by putting $x < y$ if the interval representing $x$ lies entirely to the left of the interval representing $y$.

Loosely, a random semiorder is obtained by placing $n$ unit intervals uniformly at random on an interval of given length, with the order as above. We give full details later.

Semiorders have a very special and well-understood structure; an alternative definition is that a semiorder is a partial order not containing either of the two four-element partial orders $H$ and $L$ shown in Figure 3.1 as an induced suborder. See Fishburn [14] for a proof of this and much more information about semiorders.
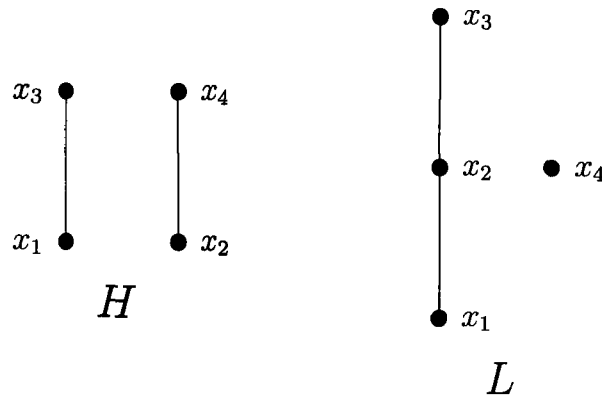
Figure 3.1: Forbidden induced suborders

As explained in Chapter 1, transitive percolation is a one-parameter family of models from the larger family of classical sequential growth models. Recall that a particular classical sequential growth model is specified by a sequence $\mathbf{t} = (t_0, t_1, \dots)$ of non-negative constants. We start with the partial order $P_0$ with one element labelled 0. At stage $n = 1, 2, \dots$, the element $n$ is added to $P_{n-1}$ and placed above all elements in $D_n$, where $D_n$ is a random subset of $\{0, 1, \dots, n-1\}$, the probability that $D_n$ is equal to a set $D$ being proportional to $t_{|D|}$. The transitive closure is taken to form the partial order $P_n$. We write $CSG(\mathbf{t})$ for the model specified by sequence $\mathbf{t}$. These models are of particular interest as they are the only ones satisfying some natural-looking conditions for discrete models of space-time—recall the conditions of "discrete general covariance" and "Bell causality" explained in Chapter 1.

It is natural to ask whether continuum limits exist for sequences of classical sequential growth models other than a sequence of transitive percolation models, and in particular whether one can obtain continuum limits that are radically different from random semiorders. In Section 3.2, we show that this is not possible.

We say that a partially ordered measure space $P$ is an *almost-semiorder* if the probability that a random sample of four elements from $P$ is isomorphic to either $H$ or $L$ is zero.

**Theorem 3.5.** *If a sequence of classical sequential growth models $(P_n)_{n=0}^{\infty}$ has a*

*continuum limit, then this limit is an almost-semiorder.*

It has been asked [24, 25] whether classical sequential growth models can be constructed to resemble a "sprinkling" from Minkowski space $M^d$, for any dimension $d \geq 2$, i.e., a partial order obtained from $M^d$ by taking points according to a Poisson process with fixed density $\lambda$. Alternatively, can a classical sequential growth model have a continuum limit resembling $M^d$? The results here demonstrate that this is not possible. Indeed, an interval $[a, b]$ of $M^d$ is a long way from being a semiorder, so no classical sequential growth model can have a region of $M^d$ as a continuum limit.

Before proving Theorems 3.4 and 3.5 we give a few observations about the probabilities $\lambda(Q; P)$.

**Lemma 3.6.** *For $Q, P$ finite labelled partial orders with $|Q| = j$, and for $k$ with $j \leq k \leq |P|$,*

$$\sum_{\substack{|Q'|=k \\ Q'|_{\{x_1,\dots,x_j\}}=Q}} \lambda(Q'; P) = \lambda(Q, P).$$

**Proof.** Fix $Q$ with $|Q| = j$. For any $k$ with $j \leq k \leq |P|$, construct a random labelled partial order by taking a random sample $x_1 \dots, x_k$ of $k$ elements from $P$. The probability that the labelled subposet on $x_1, \dots, x_j$ is equal to $Q$ is the sum of $\lambda(Q', P)$ over all labelled partial orders $Q'$ that, when restricted to $\{x_1, \dots, x_j\}$, are equal to $Q$. But this probability must be equal to $\lambda(Q; P)$, as we are only looking at the structure of the first $j$ elements sampled. $\square$

**Corollary 3.7.** *If $Q$ is a (labelled) subposet of $Q'$ then $\lambda(Q'; P) \leq \lambda(Q; P)$, for all $P$ with $|P| \geq |Q'|$.*

**Proof.** This follows immediately from Lemma 3.6. $\square$

Write $A_k$ for the $k$-element labelled antichain and $C_k$ for the $k$-element labelled chain, $\{x_1 < x_2 < \cdots < x_k\}$. We have the following result which will allow us just to consider $\lambda(C_2; P_n)$ and $\lambda(A_2; P_n)$.

**Proposition 3.8.**

(i) If $\mathbb{E}\lambda(A_2; P_n) \to 0$ as $n \to \infty$, then $\mathbb{E}\lambda(Q; P_n) \to 0$ as $n \to \infty$ for all finite labelled partial orders $Q$ that are not a chain, and $\mathbb{E}\lambda(C_k; P_n) \to 1/k!$ as $n \to \infty$ for all $k \geq 2$.

(ii) If $\mathbb{E}\lambda(C_2; P_n) \to 0$ as $n \to \infty$, then $\mathbb{E}\lambda(Q; P_n) \to 0$ as $n \to \infty$ for all finite labelled partial orders $Q$ that are not an antichain, and $\mathbb{E}\lambda(A_k; P_n) \to 1$ as $n \to \infty$ for all $k \geq 2$.

**Proof.** We show part (i). Part (ii) can be proved in a similar way.

Assume $\mathbb{E}\lambda(A_2; P_n) \to 0$ as $n \to \infty$. Fix $k \geq 2$, and let $Q$ be any labelled partial order of size $k$, but not equal to $C_k$. Define $Q'$ as a relabelled copy of $Q$ with the elements $x_1, x_2$ incomparable, which is possible since $Q \neq C_k$. Note that $\lambda(Q'; P_n) = \lambda(Q; P_n)$. Since $A_2$ is a subposet of $Q'$, we can apply Corollary 3.7 giving $\lambda(Q; P_n) \leq \lambda(A_2; P_n)$. So, $\mathbb{E}\lambda(A_2; P_n) \to 0$ as $n \to \infty$ implies that $\mathbb{E}\lambda(Q; P_n) \to 0$ as $n \to \infty$ for all $Q$ of size $k$ not equal to $C_k$. But $\sum_{|Q|=k} \lambda(Q; P_n) = 1$, and there are $k!$ labellings of the $k$-element chain, so we have $\mathbb{E}\lambda(C_k; P_n) \to 1/k!$ as $n \to \infty$.  $\square$

## 3.1  Random graph orders

We recall the definition of a random graph order.

**Definition 3.9.** Let $P_{n,p}$ be a random partial order on $[n-1] \equiv \{0, 1, \ldots, n-1\}$, formed by introducing the relation $(i, j)$ with probability $p$, independently for each pair of elements $i < j$, and then taking the transitive closure. The partial order $P_{n,p}$ is called a *random graph order*.

Note that the description of $P_{n,p}$ above is equivalent to that given earlier. In future, we will use the term random graph order, rather than transitive percolation, but the reader should be aware that the terms are essentially interchangeable.

### 3.1.1 Some results on $P_{n,p}$

We include some results of Pittel and Tungol, from [23], which we will need in order to prove the existence of a continuum limit. Results of a similar type can be found in Bollobás and Brightwell [8], and Simon, Crippa and Collenberg [27]. We change the notation slightly, for ease of use here. The following results apply to a random graph order $P_{N,\pi}$, and we will apply them with particular values for $N$ and $\pi$. Very crudely, these results can be interpreted as saying that, if $i$ and $j$ are elements of $[N-1]$, then

(i) for $\alpha > 1$, most pairs $(i,j)$ with $j - i \geq \alpha\pi^{-1}\log\pi^{-1}$ are comparable in $P_{N,\pi}$,

(ii) for $\alpha < 1$, few pairs $(i,j)$ with $0 < j - i \leq \alpha\pi^{-1}\log\pi^{-1}$ are comparable in $P_{N,\pi}$.

**Theorem 3.10** (Pittel and Tungol, [23, Theorem 4.1(3)]). *Let $X$ be the number of comparable pairs $i < j$ in $P_{N,\pi}$. Let $\pi = \alpha\log N/N$ with $\alpha > 1$. Then*

$$\mathbb{E}X = (1 + o(1))\frac{1}{2}\left(N\left(1 - \frac{1}{\alpha}\right)\right)^2.$$

Define $\gamma_N^*(0)$ to be the size of the up-set of 0 in $P_{N,\pi}$.

**Theorem 3.11** (Pittel and Tungol, [23, Theorem 2.3(1)]). *Let $\pi = \alpha\log N/N$. Suppose that $\alpha \geq 1$. If $M$ is such that*

$$f(M) \equiv \left(M - N\left(1 - \frac{1}{\alpha}\right)\right)\frac{\alpha\log N}{N} - \log\log N = O(\log\log N),$$

*then*

$$\mathbb{P}(\gamma_N^*(0) > M) = (1 + o(1))\exp\left(-\frac{1}{\alpha}e^{f(M)}\right).$$

**Theorem 3.12** (Pittel and Tungol, [23, Corollary 2.4(3)]). *Let $\pi = \alpha\log N/N$. If $\alpha = \alpha(N) < 1$ and*

$$(1 - \alpha)\log N - \log\log N \geq -2\log\log\log N,$$

*then*

$$\mathbb{E}(\gamma_N^*(0)) = (1 + o(1))N^\alpha.$$

## 3.1.2 The continuum limits of $P_{n,p}$

We show that, for suitable functions $p(n)$, the continuum limit is the semiorder defined below.

**Definition 3.13.** For $0 \leq c \leq 1$, let $S_c$ be the partially ordered measure space $([0,1], \mathcal{B}, \mu_L, \prec)$, where $\mathcal{B}$ is the family of Borel sets on $[0,1]$, the measure $\mu_L$ is the Lebesgue measure on $[0,1]$, and $\prec$ is defined by $x \prec y$ if and only if $y - x > c$.

In particular, $S_0$ is the partially ordered measure space $([0,1], \mathcal{B}, \mu_L, \prec)$ with $x \prec y$ for all $x < y$, so that $([0,1], \prec)$ is a chain, and $S_1$ is the partially ordered measure space $([0,1], \mathcal{B}, \mu_L, \prec)$ with $x \nprec y$ for all $x, y$, so that $([0,1], \prec)$ is an antichain.

By associating the number $y$ with an interval of length $c$ with left-endpoint $y$, we see immediately that $S_c$ is a semiorder. We now prove that, for certain $p(n)$, the semiorder $S_c$ is the continuum limit of our sequence of random graph orders.

**Theorem 3.14.** *The sequence of models* $(\mathcal{P}_{n,p})_{n=1}^{\infty}$ *of random graph orders has a continuum limit for* $p = p(n)$ *when either*

*(i)* $\lim_{n \to \infty} (p^{-1} \log p^{-1}/n) = 0,$

*(ii)* $\lim_{n \to \infty} (p^{-1} \log p^{-1}/n) = c$ *for some* $0 < c < 1,$ *or*

*(iii)* $\liminf_{n \to \infty} (p^{-1} \log p^{-1}/n) \geq 1.$

*The continuum limit in each case is*

*(i)* $S_0$, *i.e., a chain,*

*(ii)* $S_c$,

*(iii)* $S_1$, *i.e., an antichain.*

**Proof.** Suppose that $\lim_{n \to \infty} (p^{-1} \log p^{-1}/n) = 0$. We will show that the continuum limit is $S_0 = ([0,1], \mathcal{B}, \mu_L, \prec)$ with $x \prec y$ for all $x < y$. Since $\lambda(Q; S_0) = 0$ for all $Q$

not a chain, and $\lambda(C_k; S_0) = 1/k!$ for all $k$, by Proposition 3.8, it is enough to show that $\mathbb{E}\lambda(A_2; P_n) \to 0$ as $n \to \infty$.

Fix $\varepsilon$ with $0 < \varepsilon < 0.01$, and let $n_0$ be such that $p \geq (1/\varepsilon) \log n/n$ for all $n \geq n_0$. We can apply Theorem 3.10 with $N = n$, $\pi = (1/\varepsilon) \log N/N$, so that $\alpha = 1/\varepsilon$. We have $\mathbb{E}\lambda(A_2; P_{N,\pi}) = 1 - \mathbb{E}X/\binom{N}{2}$ which by Theorem 3.10 gives

$$\mathbb{E}\lambda(A_2; P_{n,p}) \leq \mathbb{E}\lambda(A_2; P_{n,(1/\varepsilon)\log n/n}) = 1 - \frac{(1+o(1))\frac{1}{2}(n(1-\varepsilon))^2}{\binom{n}{2}} \leq 2\varepsilon + o(1).$$

So, $\mathbb{E}\lambda(A_2; P_{n,p}) \to 0$ as required.

Now, suppose that $\liminf_{n\to\infty} (p^{-1}\log p^{-1}/n) \geq 1$. We will show that the continuum limit is $S_1 = ([0,1], \mathcal{B}, \mu_L, \prec)$ with $x \not\prec y$ for all $x, y$. Since $\lambda(Q; S_1) = 0$ for all $Q$ not an antichain, and $\lambda(A_k; S_1) = 1$ for all $k$, by Proposition 3.8, it is enough to show that $\mathbb{E}\lambda(C_2; P_n) \to 0$ as $n \to \infty$.

Fix $\varepsilon$ with $0 < \varepsilon < 0.01$. Choose $n_0$ such that $p \leq (1 + \varepsilon) \log n/n$ for $n \geq n_0$. We can apply Theorem 3.10 with $N = n$, $\pi = (1 + \varepsilon) \log N/N$, so that $\alpha = 1 + \varepsilon$. We have $\mathbb{E}\lambda(C_2; P_{N,\pi}) = \mathbb{E}X/\binom{N}{2}$ which by Theorem 3.10 gives

$$\mathbb{E}\lambda(C_2; P_{n,p}) \leq \mathbb{E}\lambda(C_2; P_{n,(1+\varepsilon)\log n/n}) = \frac{(1+o(1))\frac{1}{2}(n(1-1/(1+\varepsilon)))^2}{\binom{n}{2}} \leq \varepsilon^2 + o(1).$$

So, $\mathbb{E}\lambda(C_2; P_{n,p}) \to 0$ as required.

Finally, suppose that $\lim_{n\to\infty} (p^{-1}\log p^{-1}/n) = c$ for some $0 < c < 1$. We will show that the continuum limit is $S_c = ([0,1], \mathcal{B}, \mu_L, \prec)$ with $x \prec y$ if and only if $y - x > c$.

Fix $\varepsilon$ with $0 < \varepsilon < \min\{c, 1 - c\}$. Since $\lim_{n\to\infty} p^{-1}\log p^{-1}/n = c$ we must also have $\lim_{n\to\infty} p^{-1}\log n/n = c$, and since $c < 1$, we have $p > \log n/n$, for sufficiently large $n$. Furthermore, since

$$(1 + \varepsilon/2c)\frac{\log(c+\varepsilon)n}{(c+\varepsilon)n} = \left(\frac{c+\varepsilon/2}{c+\varepsilon}\right)\frac{\log(c+\varepsilon)n}{cn} < (1 - \delta)\frac{\log n}{cn},$$

for some $\delta > 0$, we have $p > (1 + \varepsilon/2c) \log(c + \varepsilon)n/(c + \varepsilon)n$ for sufficiently large $n$. Similarly, we have $p < (1 - \varepsilon/2c) \log(c - \varepsilon)n/(c - \varepsilon)n$ for sufficiently large

$n$. Let $n_0$ be such that $p > \log n/n$, $(1 + \varepsilon/2c) \log(c + \varepsilon)n/(c + \varepsilon)n < p < (1 - \varepsilon/2c) \log(c - \varepsilon)n/(c - \varepsilon)n$, and $n > 1/\varepsilon$ for all $n \geq n_0$.

We proceed as follows. For each $n \geq n_0$, take a random order $P_{n,p}$ according to $\mathcal{P}_{n,p}$. Define an order $\prec_n$ on $[0, 1]$, by dividing $[0, 1]$ into $n$ intervals of length $1/n$, identifying $[i/n, (i + 1)/n)$ with $i \in [n - 1]$, and putting $[i/n, (i + 1)/n)$ below $[j/n, (j + 1)/n)$ if and only if $i < j$ in $P_{n,p}$. Now for any sample of elements from $[0, 1]$ of fixed size $k$, we need that

$$\mathbb{P}(\prec_n \text{ induces different partial order from } \prec) \to 0$$

as $n \to \infty$. This is enough to prove that $\mathbb{E}\lambda(Q; P_{n,p}) \to \lambda(Q; S_c)$ as $n \to \infty$ for all finite partial orders $Q$, as follows. Let $\bar{P}_n$ be the atomless partially ordered measure space $([0, 1], \mathcal{B}, \mu_L, \prec_n)$, and suppose $Q$ is any finite partial order with $|Q| = k$. By the definitions of $\lambda(Q; P_{n,p})$ and $\lambda(Q; \bar{P}_n)$, the difference $\mathbb{E}\lambda(Q; P_{n,p}) - \mathbb{E}\lambda(Q; \bar{P}_n)$ is non-zero only because of the positive probability that in a random sample of $k$ elements from $\bar{P}_n$ some of the elements are in the same interval $[i/n, (i + 1)/n)$, for some $i$. Since the measure of these intervals tends to zero as $n \to \infty$, we have that $\mathbb{E}\lambda(Q; P_{n,p}) - \mathbb{E}\lambda(Q; \bar{P}_n) \to 0$ as $n \to \infty$. So, it is enough to show that $\mathbb{E}\lambda(Q; \bar{P}_n) \to \lambda(Q; S_c)$, which follows if $\mathbb{P}(\prec_n \text{ induces different partial order from } \prec) \to 0$ as $n \to \infty$. Indeed, it is enough to consider two elements $x, y$ chosen uniformly at random from $[0, 1]$ and show that

$$\mathbb{P}(\prec_n \text{ induces different partial order from } \prec \text{ on } \{x, y\}) \to 0 \qquad (3.1)$$

as $n \to \infty$, since for any sample $S$ of $k$ elements from $[0, 1]$,

$$\mathbb{P}(\prec_n \text{ induces different partial order from } \prec \text{ on } S) \leq \binom{k}{2}q,$$

where $q = \mathbb{P}(\prec_n \text{ induces different partial order from } \prec \text{ on } \{x, y\})$.

Call a pair of intervals $[i/n, (i + 1)/n)$ and $[j/n, (j + 1)/n)$ *good* if either

(i) $\dfrac{|i - j| - 1}{n} > c$ and $i, j$ are comparable in $P_{n,p}$, or

(ii) $\dfrac{|i-j|+1}{n} < c$ and $i,j$ are incomparable in $P_{n,p}$,

and call a pair of intervals *bad* otherwise.

We will show that the expected number of bad pairs of intervals is a small fraction of $n^2$. This will prove (3.1), since $\prec_n$ and $\prec$ will only induce different partial orders on $\{x,y\}$ if the intervals that contain $x$ and $y$ are a bad pair of intervals.

We can be rather crude with our calculations, and can afford to assume that pairs of intervals that are "too close to call" are all bad. That is, we assume that all pairs $(i,j)$ with $c - \varepsilon \le |i-j|/n \le c + \varepsilon$ are bad. There are at most $2\varepsilon n^2$ of these. For all other pairs of intervals, either $|i-j|/n > c + \varepsilon$ or $|i-j|/n < c - \varepsilon$, and we will show that almost all such pairs are good pairs.

First consider $i < j$ with $(j-i)/n > c + \varepsilon$. Such a pair $i,j$ is bad if $i,j$ are incomparable in $P_{n,p}$. So the number of bad pairs of this type is equal to the number of bad pairs of elements in $P_{n,p}$:

$$|\{(x,y) \in P_{n,p} : x,y \text{ incomparable}, y - x > (c+\varepsilon)n\}|.$$

Define an element $x < (1 - c - \varepsilon)n$ in $P_{n,p}$ to be an $\varepsilon$-*bad* element if $|U[x] \cap [x + (c+\varepsilon)n]| < \varepsilon n/2$, and an $\varepsilon$-*good* element otherwise. We will show that the number of $\varepsilon$-bad elements is small, and the number of pairs $x,y$ with $x$ an $\varepsilon$-good element and $y - x > (c + \varepsilon)n$ is also small.

We can calculate the expected number of $\varepsilon$-bad elements as follows. Let $\pi = (1+\varepsilon/2c)\log(c+\varepsilon)n/(c+\varepsilon)n$. Since $p > \pi$, the expected number of $\varepsilon$-good elements in $P_{n,p}$ is greater than the expected number of $\varepsilon$-good elements in $P_{n,\pi}$. So, working with $P_{n,\pi}$, note also that the size $|U[x] \cap [x+(c+\varepsilon)n]|$ is equivalent to $\gamma^*_{(c+\varepsilon)n}(0)$, i.e., the size of the up-set of 0 in $P_{N,\pi}$ where $N = (c+\varepsilon)n$. We want to apply Theorem 3.11 with $N = (c+\varepsilon)n$, $\pi = (1+\varepsilon/2c)\log(c+\varepsilon)n/(c+\varepsilon)n$, so $\alpha = 1 + \varepsilon/2c$. We set $M = N(1 - 1/\alpha)$, so that $f(M) = -\log\log N$ is $O(\log\log N)$ as required and

the theorem implies that

$$\mathbb{P}(\gamma^*_{(c+\varepsilon)n}(0) > M) = (1 + o(1)) \exp\left(-\frac{1}{1+\varepsilon/2c}e^{-\log\log(c+\varepsilon)n}\right)$$
$$\geq (1 + o(1))\left(1 - \frac{c}{c+\varepsilon/2}\frac{1}{\log(c+\varepsilon)n}\right).$$

Since

$$M = N(1 - 1/\alpha) = (c+\varepsilon)n\left(1 - \frac{1}{1+\varepsilon/2c}\right) = \frac{c+\varepsilon}{c+\varepsilon/2}\frac{\varepsilon n}{2} > \varepsilon n/2,$$

we have $\mathbb{P}(x$ is $\varepsilon$-bad in $P_{n,\pi}) \leq \mathbb{P}(\gamma^*_{(c+\varepsilon)n}(0) \leq M)$. Therefore, the probability that $x$ is an $\varepsilon$-bad element in $P_{n,p}$ is $O(1/\log n) + o(1)$. So, the expected number of $\varepsilon$-bad elements is $o(n)$ and assuming the worst case, that every pair of elements $(x, y)$ with $y - x > (c+\varepsilon)n$, where $x$ is $\varepsilon$-bad, is a bad pair, this gives $o(n^2)$ bad pairs.

We now need to count the number of bad pairs $(x, y)$ with $y - x > (c+\varepsilon)n$ where $x$ is $\varepsilon$-good. So $|U[x] \cap [x + (c+\varepsilon)n]| \geq \varepsilon n/2$, and the probability that $(x, y)$ is a bad pair is the probability that there are no edges between $y$ and the elements in $U[x] \cap [y]$. But $|U[x] \cap [y]| \geq |U[x] \cap [x + (c+\varepsilon)n]| \geq \varepsilon n/2$. Therefore, for $y - x > (c+\varepsilon)n$,

$$\mathbb{P}((x,y) \text{ is bad}|x \text{ is } \varepsilon\text{-good}) \leq (1-p)^{\varepsilon n/2} \leq e^{-p\varepsilon n/2} \leq n^{-\varepsilon/2}$$

Therefore the number of bad pairs $(x, y)$ with $y - x > (c+\varepsilon)n$ where $x$ is $\varepsilon$-good is $o(n^2)$.

Finally we need to count the number of pairs $i < j$ with $(j - i)/n < c - \varepsilon$ and $i, j$ comparable in $P_{n,p}$. Let $\pi = (1 - \varepsilon/2c)\log(c-\varepsilon)n/(c-\varepsilon)n$. Since $p < \pi$ the expected size $|U[x] \cap [x + (c-\varepsilon)n]|$ in $P_{n,p}$ is less than the expected size $|U[x] \cap [x + (c-\varepsilon)n]|$ in $P_{n,\pi}$. So, working with $P_{n,\pi}$, note that $|U[x] \cap [x + (c-\varepsilon)n]|$ is equivalent to $\gamma^*_{(c-\varepsilon)n}(0)$, i.e., the size of the up-set of 0 in $P_{N,\pi}$, where $N = (c-\varepsilon)n$. So, the expected number of pairs $(x, y)$ in $P_{n,p}$ with $0 < y - x < (c - \varepsilon)n$ is at most $n\mathbb{E}\gamma^*_{(c-\varepsilon)n}(0)$ which by Theorem 3.12 is $n(1 + o(1))((c - \varepsilon)n)^{1-\varepsilon/2c} = o(n^2)$.

Therefore the total number of bad pairs of intervals is at most $2\varepsilon n^2 + o(n^2)$.

Therefore, there exists $n_1 \geq n_0$ such that

$$\mathbb{P}(\prec_n \text{ induces different partial order from } \prec \text{ on } \{x,y\}) \leq 5\varepsilon$$

for all $n \geq n_1$. Since $\varepsilon$ is arbitrary we have the result. $\square$

To complete the proof of Theorem 3.4, we now show that, for all other $p(n)$, the sequence $(\mathcal{P}_{n,p})_{n=1}^{\infty}$ does not have a continuum limit. We first make the following observations.

The probability that two elements selected at random from $S_c$ are incomparable is

$$\lambda(A_2; S_c) = 1 - (1-c)^2 = 2c - c^2$$

which is monotonic in $c$ for $0 \leq c \leq 1$. So, we have

**Lemma 3.15.** *For $0 \leq c_1 \neq c_2 \leq 1$, $\lambda(A_2; S_{c_1}) \neq \lambda(A_2; S_{c_2})$.* $\square$

The following Lemma is an obvious extension to Theorem 3.14 and is stated without proof.

**Lemma 3.16.** *If we have a subsequence $(\mathcal{P}_{a_n,p})_{n=1}^{\infty}$ of random graph orders, with $p = p(a_n)$ satisfying one of conditions (i),(ii) or (iii) of Theorem 3.14, then the subsequence has a continuum limit as described in Theorem 3.14.* $\square$

**Theorem 3.17.** *If a sequence $(\mathcal{P}_n)_{n=1}^{\infty}$ of models of random graph orders has a continuum limit, then $p = p(n)$ satisfies one of conditions (i), (ii) or (iii) of Theorem 3.14.*

**Proof.** Suppose $(\mathcal{P}_{n,p})_{n=1}^{\infty}$ is a sequence of models of random graph orders with $p = p(n)$ not satisfying any of (i), (ii) or (iii). This means that

$$\liminf_{n \to \infty} p^{-1} \log p^{-1}/n < 1, \text{ and } \liminf_{n \to \infty} p^{-1} \log p^{-1}/n < \limsup_{n \to \infty} p^{-1} \log p^{-1}/n.$$

So, there exist subsequences $(a_n)$, $(b_n)$ with $\lim_{n \to \infty} p^{-1} \log p^{-1}/a_n = c_1 < 1$, where $p = p(a_n)$, and $\lim_{n \to \infty} p^{-1} \log p^{-1}/b_n = c_2 > c_1$, where $p = p(b_n)$.

So, by Lemma 3.16 the subsequence $(\mathcal{P}_{a_n,p})_{n=1}^{\infty}$ has continuum limit $S_{c_1}$ and the subsequence $(\mathcal{P}_{b_n,p})_{n=1}^{\infty}$ either has continuum limit $S_{c_2}$ or $S_1$ depending on whether $c_2 < 1$ or $c_2 \geq 1$. In either case, by Lemma 3.15 we have $\lim_{n\to\infty} \mathbb{E}\lambda(A_2; P_{a_n,p}) \neq \lim_{n\to\infty} \mathbb{E}\lambda(A_2; P_{b_n,p})$. This implies that $(\mathbb{E}\lambda(A_2; P_{n,p}))_{n=1}^{\infty}$ does not converge, and therefore $(\mathcal{P}_{n,p})_{n=1}^{\infty}$ is not compatible and so has no continuum limit.    □

This establishes Theorem 3.4.

## 3.2   Possible continuum limits of classical sequential growth models

In Section 3.1.2 we showed that the random graph order $P_{n,p}$ has a continuum limit for suitable functions $p = p(n)$ and, when it exists, the continuum limit must be the semiorder $S_c$, where $0 \leq c \leq 1$ depends on $p$. As explained earlier, random graph orders are a particular class of models from the larger family of classical sequential growth models. In this section, we show that for any sequence of classical sequential growth models, if the sequence has a continuum limit, then this limit must be an almost-semiorder, as defined earlier. Unfortunately, we do not have a complete result, like that of Theorem 3.4 for random graph orders; we provide some necessary conditions for a sequence of classical sequential growth models to have a continuum limit. The question of whether our necessary conditions are also sufficient is similar to the question of whether compatibility is a sufficient condition, and as mentioned earlier, we do not answer these questions here.

The results we give apply to all sequences t, but they are most interesting when the $t_i$ tend to zero at least exponentially quickly but, in some sense, not much more quickly. By this we mean that the general case should mirror the situation in the case of random graph orders, so that the "interesting" continuum limits occur for sequences t that are delicately balanced. We do not wish to spend time making

these statements rigorous, but to help the reader understand this point, we give the following rather loose argument. In the case where a classical sequential growth model is specified by a sequence where the $t_i$ do not tend to zero quickly enough, the growth model will produce a partial order typically denser than that produced by some random graph order, $P_{n,p}$, satisfying condition (i) of Theorem 3.14. Therefore, we would expect the continuum limit of the growth model to be denser than the continuum limit of $P_{n,p}$, which, by Theorem 3.14, is a chain. Hence, we expect the continuum limit of the growth model to be a chain. On the other hand, if a classical sequential growth model is specified by a sequence where the $t_i$ tend to zero too quickly, the growth model will produce a partial order typically sparser than that produced by some random graph order satisfying condition (iii) of Theorem 3.14, and therefore we would expect the continuum limit of the growth model to be sparser than that of the random graph order, and hence an antichain.

To give a specific example, we note that the continuum limit of a sequence of random binary growth models (or indeed, any models where $t_i = 0$ for all $i$ greater than some constant independent of $n$) is an antichain. This can be seen by noting that for any $r > \varepsilon n$, the expected size of the up-set $U[r] \cap [n]$ is bounded by a constant (dependent on $\varepsilon$, but not $n$), which follows from a simple path-counting argument as in Lemma 2.13. This implies that, for any $\varepsilon > 0$, the expected number of comparable pairs of elements is less than $2\varepsilon n^2$ for sufficiently large $n$ and by Proposition 3.8 this is enough to show that the continuum limit is an antichain.

Our task is to show that, for any continuum limit $P_\infty$ of a sequence of classical sequential growth models, $\lambda(H; P_\infty) = \lambda(L; P_\infty) = 0$. Informally, we need to show that, in any classical sequential growth model $CSG(\mathbf{t})$, the number of copies of $H$ and $L$ as subposets of $CSG(\mathbf{t})$ is small. To do this, we show that $CSG(\mathbf{t})$ has a threshold "level" such that there are very few comparable pairs below the threshold, whereas above the threshold, where the number of comparable pairs may become significant, the model behaves very roughly like a random graph order, with every

new element selecting a significant proportion of the existing elements.

To this end, we present some lemmas describing some properties of classical sequential growth models. We believe these results to be important in their own right, since they give particularly qualitative descriptions of a model $CSG(t)$ without referring to the sequence $t$ that specifies the model.

**Recall** that $D_x$ is the set of elements selected by element $x$, and $U[x]$ is the up-set of $x$. Note that $D_x$ is not the same as $D[x]$, the down-set of $x$. We begin with the following observation on the expected size of $D_x$.

**Lemma 3.18.** *For any classical sequential growth model, $\mathbb{E}(|D_x|)$ is increasing in $x$.*

**Proof.** We show that for any $x$, we have the inequality $\mathbb{E}(|D_x|) \le \mathbb{E}(|D_{x+1}|)$.

Suppose the classical sequential growth model is defined by the sequence $t = (t_0, t_1, \dots)$. Note that

$$\mathbb{E}(|D_x|) = \sum_{j=0}^{x} j \mathbb{P}(|D_x| = j) = \frac{\sum_{j=0}^{x} j \binom{x}{j} t_j}{\sum_{j=0}^{x} \binom{x}{j} t_j}$$

depends only on $t_0, t_1, \dots, t_x$, and similarly

$$\mathbb{E}(|D_{x+1}|) = \frac{\sum_{j=0}^{x+1} j \binom{x+1}{j} t_j}{\sum_{j=0}^{x+1} \binom{x+1}{j} t_j}$$

depends on $t_0, t_1, \dots, t_{x+1}$.

Note that, for fixed $t_0, t_1, \dots, t_x$ the probability $\mathbb{P}(|D_{x+1}| = x + 1)$ is increasing in $t_{x+1}$ and all other probabilities $\mathbb{P}(|D_{x+1}| = j)$ are decreasing in $t_{x+1}$. This means that $\mathbb{E}(|D_{x+1}|)$ is increasing in $t_{x+1}$ and we have

$$\mathbb{E}(|D_{x+1}|) \ge \frac{\sum_{j=0}^{x} j \binom{x+1}{j} t_j}{\sum_{j=0}^{x} \binom{x+1}{j} t_j}. \tag{3.2}$$

Now, note that $\binom{x+1}{j} = \frac{x+1}{x+1-j} \binom{x}{j}$ so the inequality (3.2) becomes

$$\mathbb{E}(|D_{x+1}|) \ge \frac{\sum_{j=0}^{x} \frac{j}{x+1-j} \binom{x}{j} t_j}{\sum_{j=0}^{x} \frac{1}{x+1-j} \binom{x}{j} t_j}$$

and it remains to prove that

$$\frac{\sum_{j=0}^{x} \frac{j}{x+1-j} \binom{x}{j} t_j}{\sum_{j=0}^{x} \frac{1}{x+1-j} \binom{x}{j} t_j} \geq \frac{\sum_{j=0}^{x} j \binom{x}{j} t_j}{\sum_{j=0}^{x} \binom{x}{j} t_j}$$

which follows from Chebyshev's Sum Inequality (see, e.g., [16, Theorem 43]), since both $j$ and $1/(x+1-j)$ are increasing on $\{0, 1, \ldots, x\}$.   □

For a fixed $CSG(\mathbf{t})$, consider the process up to stage $y_0$, which produces a partial order on the ground set $[y_0]$. Informally, the following lemma says that, if the expected sizes $\mathbb{E}(|D_y|)$ are small (these depend only on $\mathbf{t}$ and not the partial order produced), then apart from the first $\varepsilon n$ elements the expected sizes of the up-sets of the elements in $[y_0]$ are also small.

**Lemma 3.19.** *For $0 < \varepsilon < 1$, $\delta > 0$ and $n \in \mathbb{N}$, if $\mathbb{E}(|D_{y_0}|) < \delta \log n$ for some $y_0 \leq n - 1$ then $\mathbb{E}(|U[x] \cap [y_0]|) \leq n^{\delta \log (1/\varepsilon)}$ for all $x \in [\varepsilon n, n - 1]$.*

**Proof.** Note that by Lemma 3.18 we have that $\mathbb{E}(|D_y|) < \delta \log n$ for all $y \leq y_0$. We use the fact that

$$\mathbb{P}(y \text{ selects } x) = \frac{\mathbb{E}(|D_y|)}{y} < \frac{\delta \log n}{y} \tag{3.3}$$

for all $x < y \leq y_0$, and count $\mathbb{E}(|U[x] \cap [y_0]|)$ by a path-counting method.

Define a path from $x$ to $y$ to be a sequence of elements $s_0 < s_1 < \cdots < s_k$, with $s_0 = x, s_k = y$, such that $s_j$ selects $s_{j-1}$ for all $j = 1, \ldots, k$. So, the probability that any given sequence $s_0 < s_1 < \cdots < s_k$ is a path is exactly

$$\prod_{j=2}^{k} \mathbb{P}(s_j \text{ selects } s_{j-1})$$

and the expected number of paths from $x$ to some $y$ with $y \leq y_0$ is

$$\cdot \sum_{s_1 < s_2 < \cdots < s_k \in [x+1, y_0]} \mathbb{P}(s_1 \text{ selects } x) \prod_{j=2}^{k} \mathbb{P}(s_j \text{ selects } s_{j-1}).$$

Since every element in $U[x] \cap [y_0]$ must be an end point of such a path, this expected

number of paths is an upper bound on $\mathbb{E}(|U[x] \cap [y_0]|)$. Using (3.3), we have

$$
\begin{aligned}
\mathbb{E}(|U[x] \cap [y_0]|) &\leq \sum_{\substack{s_1 < s_2 < \cdots < s_k \\ \in [x+1, y_0]}} \prod_{j=1}^{k} \frac{\delta \log n}{s_j} = \prod_{s=x+1}^{y_0} \left( 1 + \frac{\delta \log n}{s} \right) \\
&\leq \exp \left( \sum_{s=x+1}^{y_0} \frac{\delta \log n}{s} \right) \\
&\leq \exp \left( \delta \log n \int_x^{y_0} \frac{1}{s} ds \right) \\
&\leq \left( \frac{y_0}{x} \right)^{\delta \log n} \leq (1/\varepsilon)^{\delta \log n} = n^{\delta \log(1/\varepsilon)}
\end{aligned}
$$

where the equality in the first line is seen by expanding out the product on the right hand side and noting that each term appears in the sum on the left hand side. The final inequality comes from the fact that $y_0 \leq n$ and $x \geq \varepsilon n$.  $\square$

Using Markov's inequality with the conclusion of Lemma 3.19 gives the following result.

**Lemma 3.20.** *For* $0 < \varepsilon < 1$, $\delta = (2 \log(1/\varepsilon))^{-1}$ *and sufficiently large* $n$, *if* $\mathbb{E}(|D_{y_0}|) < \delta \log n$ *for some* $y_0 \leq n - 1$ *then* $\mathbb{P}(|U[x] \cap [y_0]| \geq \varepsilon n/2) < \varepsilon$ *for all* $x \in [\varepsilon n, n - 1]$.

**Proof.** By Lemma 3.19 we have that $\mathbb{E}(|U[x] \cap [y_0]|) \leq n^{\delta \log(1/\varepsilon)}$ for all $x \in [\varepsilon n, n-1]$. So, Markov's inequality implies that $\mathbb{P}(|U[x] \cap [y_0]| \geq \varepsilon n/2) \leq 2n^{\delta \log(1/\varepsilon)}/\varepsilon n = (2/\varepsilon)n^{-1/2}$, which gives the result.  $\square$

So, we have that for sufficiently large $n$, in order to get large up-sets, we require that the model $CSG(\mathbf{t})$ is such that the expected size $\mathbb{E}(|D_{y_0}|)$ is at least $\delta \log n$ for some $y_0$. The next lemma shows that if at any stage $y$ the expected size $\mathbb{E}(|D_y|)$ *is* this large, then we do not have to continue the growth process much further before we find $z \geq y$ having a small probability that the random variable $|D_z|$ is small.

**Lemma 3.21.** *For* $\varepsilon > 0$ *and sufficiently large* $n$, *if* $\mathbb{E}(|D_y|) \geq \delta \log n$ *for some* $y \in [\varepsilon n, n - 1 - n/\sqrt{\log n}]$ *and* $\delta > 0$, *then* $\mathbb{P}(|D_{y+n/\sqrt{\log n}}| \leq \sqrt{\log n}) < \varepsilon/8$.

**Proof.** Fix $\varepsilon > 0$. Suppose that $\mathbb{E}(|D_y|) \geq \delta \log n$, for some constant $\delta > 0$. Since

$$\mathbb{E}(|D_y|) = \frac{\sum_{i=0}^{y} i \binom{y}{i} t_i}{\sum_{i=0}^{y} \binom{y}{i} t_i},$$

we have

$$\sum_{i=0}^{y} i \binom{y}{i} t_i \geq \delta \log n \sum_{i=0}^{y} \binom{y}{i} t_i. \tag{3.4}$$

Let $j = n/\sqrt{\log n}$ and $M = \sqrt{\log n}$. We need to bound from above the probability

$$\mathbb{P}(|D_{y+j}| \leq M) = \frac{\sum_{i=0}^{M} \binom{y+j}{i} t_i}{\sum_{i=0}^{y+j} \binom{y+j}{i} t_i}. \tag{3.5}$$

We use the following upper and lower bounds for $\binom{y+j}{i} / \binom{y}{i}$. We have,

$$\frac{\binom{y+j}{i}}{\binom{y}{i}} = \frac{(y+j)(y+j-1)\cdots(y+j-i+1)}{y(y-1)\cdots(y-i+1)} \geq \left(\frac{y+j}{y}\right)^i = \left(1 + \frac{j}{y}\right)^i$$

But $j = n/\sqrt{\log n}$ and $y \geq \varepsilon n$, so $j/y \leq 1/(\varepsilon\sqrt{\log n})$ so for any $\eta > 0$ we have

$$\frac{\binom{y+j}{i}}{\binom{y}{i}} \geq (1 - \eta)e^{ij/y} \tag{3.6}$$

for all $i$, for sufficiently large $n$.

Also,

$$\frac{\binom{y+j}{i}}{\binom{y}{i}} = \frac{(y+j)(y+j-1)\cdots(y+j-i+1)}{y(y-1)\cdots(y-i+1)} \leq \left(\frac{y+j-i+1}{y-i+1}\right)^i \leq e^{ij/(y-i+1)}$$

So, for $i \leq M = \sqrt{\log n}$ we have

$$\frac{\binom{y+j}{i}}{\binom{y}{i}} \leq e^{2Mj/y} \tag{3.7}$$

for sufficiently large $n$.

So, using the upper bound for $\binom{y+j}{i} / \binom{y}{i}$ in the numerator in (3.5) and the lower bound in the denominator, we have

$$\mathbb{P}(|D_{y+j}| \leq M) \leq \frac{e^{2Mj/y} \sum_{i=0}^{M} \binom{y}{i} t_i}{(1-\eta) \sum_{i=0}^{y+j} e^{ij/y} \binom{y}{i} t_i} \leq \frac{e^{2Mj/y} \sum_{i=0}^{y} \binom{y}{i} t_i}{(1-\eta) \sum_{i=0}^{y} e^{ij/y} \binom{y}{i} t_i}$$

and using (3.4) we have

$$\mathbb{P}(|D_{y+j}| \leq M) \leq \frac{e^{2Mj/y}}{(1-\eta)\delta \log n} \frac{\sum_{i=0}^{y} i \binom{y}{i} t_i}{\sum_{i=0}^{y} e^{ij/y} \binom{y}{i} t_i}.$$

Finally, we use the fact that $i/e^{ij/y}$ is maximised when $i = y/j$ so that $i/e^{ij/y} \leq e^{-1}y/j$.

So, we have

$$\mathbb{P}(|D_{y+j}| \leq M) \leq \frac{e^{2Mj/y}e^{-1}y/j}{(1-\eta)\delta\log n}$$

which, after substituting $j = n/\sqrt{\log n}$, $M = \sqrt{\log n}$ and $\varepsilon n \leq y \leq n$, gives

$$\mathbb{P}(|D_{y+n/\sqrt{\log n}}| \leq \sqrt{\log n}) \leq \frac{e^{2/\varepsilon - 1}}{(1-\eta)\delta\sqrt{\log n}}$$

for sufficiently large $n$. Therefore

$$\mathbb{P}(|D_{y+n/\sqrt{\log n}}| \leq \sqrt{\log n}) < \varepsilon/8$$

for sufficiently large $n$, as required. $\qquad\square$

Combining the previous two results, we can now define a threshold in a classical sequential growth model. Informally, the threshold has the property that, for sufficiently large $n$, below the threshold we have very few comparable pairs, whereas just above the threshold, where we begin to get a significant number of comparable pairs, a high proportion of elements select a significant number of existing elements. We shall see that, for example, this means that the expected proportion of triples $x_1 < x_2 < x_3$ with $x_1$ and $x_2$ comparable and $x_1$ and $x_3$ incomparable is small. Importantly, the window between "below the threshold" and "just above the threshold" is small enough not to be a problem.

**Lemma 3.22.** *For any classical sequential growth model $CSG(\mathbf{t})$, any $\varepsilon$ with $0 < \varepsilon < 1$ and any $n \in \mathbb{N}$, there exists a threshold $y_0 = y_0(\mathbf{t}, \varepsilon, n)$ defined as*

*(i) if $y \leq y_0$ then $\mathbb{P}(|U[x] \cap [y]| \geq \varepsilon n/2) < \varepsilon$ for all $x \in [\varepsilon n, n-1]$,*

*(ii) if $y_0 < y \leq n-1$ then $\mathbb{P}(|U[x] \cap [y]| \geq \varepsilon n/2) \geq \varepsilon$ for some $x \in [\varepsilon n, n-1]$.*

*Furthermore, if $n$ is sufficiently large, then*

*(a) $\mathbb{E}(|D_y|) \geq \delta\log n$ for all $y_0 < y \leq n-1$, where $\delta = (2\log(1/\varepsilon))^{-1}$,*

*(b)* $\mathbb{P}(|D_y| \leq \sqrt{\log n}) < \varepsilon/8$ *for all* $y_0 + \varepsilon n/4 \leq y \leq n - 1$.

**Proof.** Fix $\mathbf{t}, \varepsilon$, and $n$. Note that for fixed $x \in [\varepsilon n, n - 1]$ the probability $\mathbb{P}(|U[x] \cap [y]| \geq \varepsilon n/2)$ is zero for $y < x - 1 + \varepsilon n/2$ and increasing in $y$, so we have a threshold as claimed. Specifically, $y_0$ is the minimum over all $x \in [\varepsilon n, n - 1]$ of the maximum $y \leq n - 1$ satisfying $\mathbb{P}(|U[x] \cap [y]| \geq \varepsilon n/2) < \varepsilon$. By the previous remark, this is well defined and gives the threshold as required.

Now, suppose $n$ is large enough so that we can apply Lemmas 3.20 and 3.21. To prove (a), we apply Lemma 3.20, which implies that $\mathbb{E}(|D_y|) \geq \delta \log n$ for all $y_0 < y \leq n - 1$, with $\delta$ as claimed. To prove (b), we apply Lemma 3.21 for each $y$ satisfying (a), which implies that $\mathbb{P}(|D_{y+n/\sqrt{\log n}}| \leq \sqrt{\log n}) < \varepsilon/8$ for all $y_0 < y \leq n - 1 - n/\sqrt{\log n}$. For sufficiently large $n$, we have $n/\sqrt{\log n} < \varepsilon n/4$, so that $\mathbb{P}(|D_y| \leq \sqrt{\log n}) < \varepsilon/8$ for all $y_0 + \varepsilon n/4 \leq y \leq n - 1$. $\qquad\square$

For a particular partial order $P_n$ arising from a growth process on $[n - 1]$, and elements $x, y \in [\varepsilon n, n - 1]$, we say that $x$ is a *y-good element (in $P_n$)* if $|U[x] \cap [y]| < \varepsilon n/2$, and a *y-bad element (in $P_n$)* otherwise. So $y_0$ as defined above is the maximum value $y$ such that, for all elements $x$ in $[\varepsilon n, n - 1]$, the probability that $x$ is $y$-bad is less than $\varepsilon$. In particular, the expected number of $y_0$-bad elements is at most $\varepsilon n$.

For any subset $A \subseteq [n - 1]$ define $\max A$ to be the largest element in $A$. We have the following lemma, which states that if a set is at least some (small) constant proportion of $[n - 1]$, then the expected number of elements above the threshold $y_0$ not selecting elements in the set is at most some (small) constant fraction of $n$.

**Lemma 3.23.** *For any $\varepsilon$ with $0 < \varepsilon < 1$, and sufficiently large $n$, given any subset $A \subseteq [n - 1]$ with $|A| > \varepsilon n/2$, and $m = \max\{y_0 + \varepsilon n/4, \max A + 1\}$, we have*

$$\mathbb{E}(\text{number of elements in } [m, n - 1] \text{ not selecting an element of } A) \leq \varepsilon n/4,$$

*where $y_0$ is the threshold defined in Lemma 3.22.*

**Proof.** By Lemma 3.22, for $y \in [m, n - 1]$ we have $\mathbb{P}(|D_y| \leq \sqrt{\log n}) < \varepsilon/8$,

for sufficiently large $n$. So, the expected number of $y \in [m, n-1]$ with $|D_y| \leq \sqrt{\log n}$ is less than $\varepsilon n/8$, for sufficiently large $n$. For each $y \in [m, n-1]$ with $|D_y| > \sqrt{\log n}$ the probability that $y$ does not select an element in $A$ is less than $(1 - \varepsilon/2)^{\sqrt{\log n}} \leq e^{-\varepsilon \sqrt{\log n}/2} < \varepsilon/8$ for sufficiently large $n$. So, in total the expected number of elements in $[m, n-1]$ not selecting an element in $A$ is less than $\varepsilon n/4$, for sufficiently large $n$. $\square$

We also require a result of Chernoff, which bounds large deviations of a binomial random variable from its mean. For further details see, for example, [2, Appendix A].

**Theorem 3.24** (Chernoff). *If $X \sim B(N, \pi)$ and $a > 0$, then $\mathbb{P}(X < (1 - a)\pi N) < e^{-a^2 \pi N/2}$.* $\square$

Using Lemmas 3.22 and 3.23 and Theorem 3.24, we now prove Theorem 3.5, which states that a continuum limit of a sequence $(\mathcal{P}_n)_{n=1}^{\infty}$ of classical sequential growth models is an almost-semiorder.

**Proof of Theorem 3.5.** Suppose $P_{\infty}$ is the continuum limit of $(\mathcal{P}_n)_{n=1}^{\infty}$. Recall that $H$ and $L$ are the four-element partial orders in Figure 3.1. We will show that both $\mathbb{E}\lambda(H; P_n) \to 0$ and $\mathbb{E}\lambda(L; P_n) \to 0$ as $n \to \infty$, where $P_n$ is a random partial order taken from $\mathcal{P}_n$, which implies that both $\lambda(H; P_{\infty}) = 0$ and $\lambda(L; P_{\infty}) = 0$.

**Claim 3.1.** $\mathbb{E}\lambda(H; P_n) \to 0$ *as* $n \to \infty$.

**Proof of Claim 3.1.** Call a quadruple $X = (x_1, x_2, x_3, x_4) \in [n-1]^{(4)}$ an $H$-*quadruple* if the partial order on $X$ induced by the order on $P_n$ is equal to the partial order $H$. Fix $\varepsilon > 0$. We will show that, for sufficiently large $n$, the expected number of $H$-quadruples is less than $5\varepsilon n^4$. This implies that $\mathbb{E}\lambda(H; P_n) < 6\varepsilon$ for sufficiently large $n$.

The number of $H$-quadruples including an element below $\varepsilon n$ is certainly at most $2\varepsilon n^4$, so we may restrict attention to quadruples all of whose entries are in $[\varepsilon n, n-1]$.

We now fix any pair of elements $(x_1, x_2)$ from $[\varepsilon n, n-1]$, and estimate the

expected number of $H$-quadruples with these elements as the first two entries, which is exactly

$$\mathbb{E}\big(|U[x_1] \setminus U[x_2]| \times |U[x_2] \setminus U[x_1]|\big).$$

This is certainly at most $n$ times the expectation of the minimum of these two sizes. We claim that, for any choice of $(x_1, x_2)$, the expected value of $\min\{|U[x_1] \setminus U[x_2]|, |U[x_2] \setminus U[x_1]|\}$ is at most $3\varepsilon n$; this will suffice to prove the claim.

Recall the threshold $y_0 = y_0(\mathrm{t}, \varepsilon, n)$, as defined in Lemma 3.22, with:

(i) if $y \le y_0$, then $\mathbb{P}(x$ is a $y$-bad element$) < \varepsilon$ for all $x \in [\varepsilon n, n-1]$,

(ii) if $y_0 < y \le n-1$, then $\mathbb{P}(x$ is a $y$-bad element$) \ge \varepsilon$ for some $x \in [\varepsilon n, n-1]$.

We consider increasing $y$ until one of the two sets $U[x_1] \cap [y]$ and $U[x_2] \cap [y]$ (without loss of generality the first) reaches size $\varepsilon n/2$. To be precise, define $y_1$ such that, for all $y \le y_1$, both $x_1$ and $x_2$ are $y$-good, and for all $y > y_1$, one of them (without loss of generality $x_1$) is $y$-bad.

One of the following events occurs.

(a) $y_1 = n-1$, as neither set ever reaches size $\varepsilon n/2$,

(b) $y_1 < y_0$,

(c) $y_0 \le y_1 < n-1$.

If event (a) occurs, then certainly $|U[x_1] \setminus U[x_2]| \le \varepsilon n/2 < \varepsilon n$.

The probability of event (b) is at most $2\varepsilon$, by definition of $y_0$ and $y_1$.

We now consider the case where event (c) occurs, and condition on the event that $y_1$ takes some particular value in the range $[y_0, n-2]$; note that this event depends only on the sets selected by those elements up to and including $y_1 + 1$; in what follows we will only consider the sets selected by elements beyond $y_1 + 1$, so we can effectively ignore the conditioning.

By definition of $y_1$, we have $|U[x_1] \cap [y_1+1]| = \lfloor \varepsilon n/2 \rfloor + 1$ and $|U[x_2] \cap [y_1]| < \varepsilon n/2$. We now show that the expected size of $U[x_2] \setminus U[x_1]$ is small.

Since the set $U[x_1] \cap [y_1+1]$ has size greater than $\varepsilon n/2$, and $y_1 \geq y_0$, Lemma 3.23 implies that the expected number of elements in $[y_1 + \varepsilon n/4, n-1]$ not selecting an element from $U[x_1] \cap [y_1+1]$ is less than $\varepsilon n/4$, for sufficiently large $n$. So the expected number of elements in $[y_1 + \varepsilon n/4, n-1]$ not above $x_1$ is less than $\varepsilon n/4$, for sufficiently large $n$.

So, conditioned on this value of $y_1$, the expected size of $U[x_2] \setminus U[x_1]$ is at most

$$|U[x_2] \cap [y_1]| + |[y_1+1, y_1+\varepsilon n/4]| + \mathbb{E}|[y_1+\varepsilon n/4, n-1] \setminus U[x_1]|$$

$$< \varepsilon n/2 + \varepsilon n/4 + \varepsilon n/4 = \varepsilon n,$$

for sufficiently large $n$.

Considering all the possible cases (a)-(c), we now have that the expected value of $\min\{|U[x_1] \setminus U[x_2]|, |U[x_2] \setminus U[x_1]|\}$ is at most $\varepsilon n + n\mathbb{P}((\text{b}) \text{ occurs}) < 3\varepsilon n$, for sufficiently large $n$. This completes the proof. $\square$

**Claim 3.2.** $\mathbb{E}\lambda(L; P_n) \to 0$ as $n \to \infty$.

**Proof of Claim 3.2.**  Call a quadruple $X = (x_1, x_2, x_3, x_4) \in [n-1]^{(4)}$ an $L$-quadruple if the partial order on $X$ induced by the order on $P_n$ is equal to the partial order $L$. Fix $\varepsilon > 0$. We will show that, for sufficiently large $n$, the expected number of $L$-quadruples is less than $6\varepsilon n^4$. This implies that $\mathbb{E}\lambda(L; P_n) < 7\varepsilon$ for sufficiently large $n$.

Generate a partial order $P_n$ on $[n-1]$ according to the classical sequential growth process defined by t, and consider any quadruple $(x_1, x_2, x_3, x_4)$ of elements from $[n-1]$, with $\varepsilon n \leq x_1 < x_2 < x_3$ (in $[n-1]$). Set $y_1 = y_1(P_n, x_1)$ to be the largest element such that $x_1$ is a $y_1$-good element. One of the following is true.

(a) $y_1 \leq y_0$,

(b) $x_2 \leq y_1 + \varepsilon n/4$,

(c) $x_2 - \varepsilon n/4 \leq x_4 \leq x_2$,

(d) $y_1 > y_0$, $x_2 > y_1 + \varepsilon n/4$ and $x_4 > x_2$,

(e) $y_1 > y_0$, $x_2 > y_1 + \varepsilon n/4$ and $x_4 < x_2 - \varepsilon n/4$.

We show that the expected number of $L$-quadruples satisfying each of the above sets of conditions is at most $\varepsilon n^4$. This gives the expected number of $L$-quadruples to be at most $6\varepsilon n^4$, including those with $x_1 < \varepsilon n$. The above claim is immediate in case (c). For case (a), the condition implies that $x_1$ is $y_0$-bad, and we know that the expected number of $y_0$-bad elements is at most $\varepsilon n$. The bound for case (b) follows immediately from the definition of $y_1$: the condition implies that $x_2$ lies in one of the small sets $U[x_1] \cap [y_1]$ or $[y_1 + 1, y_1 + \varepsilon n/4]$.

For case (d), it suffices to prove that, for fixed $x_1, x_2$ with $x_2 \geq m = \max\{y_0 + \varepsilon n/4, y_1 + 2\}$, conditioned on the growth process up to $m$, the expected number of elements $x_4 \in [x_2, n-1]$ with $x_4$ not above $x_1$ is less than $\varepsilon n/4$. We apply Lemma 3.23 to the set $U[x_1] \cap [y_1 + 1]$, which is of size $\lfloor \varepsilon n/2 \rfloor + 1$ by definition of $y_1$. Since $\max(U[x_1] \cap [y_1 + 1]) = y_1 + 1$, we have that the expectation

$\mathbb{E}$(number of elements in $[m, n-1]$ not selecting an element of $U[x_1] \cap [y_1 + 1]$)

is less than or equal to $\varepsilon n/4$, which implies the desired result.

Turning finally to case (e), we fix $x_2$ with $x_2 > y_0 + \varepsilon n/4$, and $x_4$ with $x_4 < x_2 - \varepsilon n/4$, and condition on the growth process up to $x_4$. It suffices to show that $\mathbb{E}|U[x_2] \setminus U[x_4]| < \varepsilon n$.

Since $x_2 > y_0 + \varepsilon n/4$, we have $\mathbb{E}(|D_z|) \geq \delta \log n$ for all $z \in [x_2 - \varepsilon n/4, x_2 - 1]$, where $\delta = (2 \log(1/\varepsilon))^{-1}$. So,

$$\mathbb{P}(z \text{ selects } x_4) = \frac{\mathbb{E}(|D_z|)}{z} \geq \frac{\delta \log n}{n},$$

which means that the probability that fewer than $\varepsilon\delta \log n/8$ of the elements in $[x_2 - \varepsilon n/4, x_2 - 1]$ select $x_4$ is less than

$$\mathbb{P}(X < \varepsilon\delta \log n/8)$$

where $X$ has the binomial distribution $B(\varepsilon n/4, \delta \log n/n)$. Using Theorem 3.24 with $a = 1/2$, $N = \varepsilon n/4$ and $\pi = \delta \log n/n$ gives

$$\mathbb{P}(X < \varepsilon\delta \log n/8) < e^{-\varepsilon\delta \log n/32} < \varepsilon/8$$

for sufficiently large $n$. Let $Z = U[x_4] \cap [x_2 - \varepsilon n/4, x_2 - 1]$. With probability at least $1 - \varepsilon/8$ we have $|Z| \geq \varepsilon\delta \log n/8$. We next show that, if this is the case, then, conditioned on the growth process up to $x_2$, with high probability either $|U[x_2]| < \varepsilon n/2$, or $U[x_4] \cap [y]$ reaches size $\varepsilon n/2$ before $U[x_2] \cap [y]$ does.

Recall that a *path* in a poset $P$ arising from a growth process is a sequence $a_1, a_2, \ldots, a_k$ of elements such that each element selects the previous one; we say the path has start point $a_1$ and endpoint $a_k$. For $z \leq x_2 \leq y$, define $P_z^y$ to be the set of all elements in $[y]$ that are an endpoint of some path with start point equal to $z$ and all other elements in the path in $[x_2 + 1, y]$. The starting point for this definition is that $P_z^{x_2} = \{z\}$ for each $z \leq x_2$.

Note that $P_z^y \subseteq U[z] \cap [y]$, with equality if $z = x_2$. We claim also that the processes $(P_z^y \setminus \{z\})_{y=x_2}^{n-1}$ are identically distributed, independent of the growth process up to $x_2$, for all $z \in Z \cup \{x_2\}$. To see this, note that $P_z^{x_2} \setminus \{z\} = \emptyset$ for each $z \leq x_2$, and that, for each $y > x_2$, $y$ enters $P_z^y$ if and only if it selects an element of $P_z^{y-1}$, an event whose probability depends only on $|P_z^{y-1}|$.

Now consider all the identically distributed size processes $(|P_z^y|)_{y=x_2}^{n-1}$, for $z \in Z \cup \{x_2\}$. The probability that $|P_{x_2}^y|$ reaches size $\varepsilon n/2$ before any of the other processes is, by symmetry, at most $1/(|Z| + 1) \leq 8/(\varepsilon\delta \log n) \leq \varepsilon/8$, for sufficiently large $n$. So, with total probability at least $1 - \varepsilon/4$, we have either (i) $|P_{x_2}^{n-1}| < \varepsilon n/2$, or (ii) there exists some $z \in Z$ and some $y \geq x_2$ with $|P_z^y| \geq \varepsilon n/2$ and $|P_{x_2}^y| \leq \varepsilon n/2$.

If (i) occurs, then $|U[x_2]| = |P_{x_2}^{n-1}| < \varepsilon n/2$. We will prove that, if (ii) occurs, then, conditioned on the growth process up to $y$, the expected size of $U[x_2] \setminus U[x_4]$ is less than $3\varepsilon n/4$. This will imply that the expected size of $U[x_2] \setminus U[x_4]$, conditioned only on the assumptions in (e), is less than $\varepsilon n$, as we require.

Suppose then that (ii) occurs, when we have $|U[x_4] \cap [y]| > |P_z^y| \geq \varepsilon n/2$ and $|U[x_2] \cap [y]| = |P_{x_2}^y| \leq \varepsilon n/2$. Applying Lemma 3.23 to the set $U[x_4] \cap [y]$ implies that, conditioned on the process up to $y$, the expected number of elements in $[y+1, n-1]$ not above $x_4$ is less than $\varepsilon n/4$, so the expected size of $U[x_2] \setminus U[x_4]$ is indeed less than $\varepsilon n/2 + \varepsilon n/4 = 3\varepsilon n/4$.

This completes the proof.    □

Combining the claims, we have that the continuum limit $P_\infty$ satisfies $\lambda(H; P_\infty) = 0$ and $\lambda(L; P_\infty) = 0$ so that $P_\infty$ is an almost-semiorder.    □

# Part II

# Maps of rooted trees into complete trees

This part is concerned with counting embeddings of trees into complete trees. In [17], Kubicki, Lehel and Morayne proved that for binary trees $T_1$ and $T_2$ with $T_1$ a subtree of $T_2$, the proportion of the embeddings of $T_1$ into a complete binary tree that map to the root is no more than the proportion of the embeddings of $T_2$ into the complete binary tree that map to the root. They conjectured that this inequality holds even for $T_1, T_2$ not binary. Here, we show that the conjecture is false and look at different generalisations of their result.

In Chapter 4 we give some background and motivation for this work, and introduce our notation.

In Chapter 5 we provide an algorithm for calculating the number of embeddings of a tree into a complete binary tree and the number that map to the root of the complete binary tree. Using this algorithm for a particular pair of trees we provide a counterexample to the conjecture of Kubicki, Lehel and Morayne.

In Chapter 6 we investigate the asymptotic behaviour of the number of embeddings as the height of the complete binary tree tends to infinity. Using this behaviour we are able to give conditions on when a pair of trees will be a counterexample for all large enough complete binary trees. Using this we construct a family of pairs of trees which are such "asymptotic counterexamples".

In Chapter 7 we show that the results in Chapters 5 and 6 can be reformulated for embeddings of trees into a complete $p$-ary tree and we state and prove some of these more general results.

In Chapter 8 we generalise the result of Kubicki, Lehel and Morayne to embeddings of binary trees into a complete $p$-ary (rather than into a complete binary tree). Our proof employs the FKG-inequality, a powerful result which gives correlation inequalities for events on distributive lattices. Therefore, we can view the Kubicki, Lehel and Morayne result as one of many possible correlation inequalities for embeddings of binary trees into complete $p$-ary trees. In this light we see that

the case of binary trees is special; we cannot use this distributive lattice method when we generalise to embeddings of arbitrary trees. We give an example where a correlation inequality for embeddings of binary trees does not hold if we generalise to embeddings of arbitrary trees.

However, we show that we can generalise to arbitrary trees, if we instead look at order-preserving maps from trees to complete trees. In other words, the conjectured inequality is true, if we count order-preserving maps rather than embeddings. In this case, we are able to apply the FKG-inequality to get correlation inequalities for order-preserving maps of arbitrary trees into complete trees. This is true for both strict and weak order-preserving maps. (Formal definitions can be found in Chapter 8, but the main difference between strict and weak order-preserving maps is that a strict order-preserving map cannot map two comparable elements to the same element, whereas a weak order-preserving map can.) We finish the chapter with some related open problems.

In Chapter 9 we look at some lemmas for product lattices, which give alternative sufficient conditions for applying the FKG-inequality. We show that some of the conditions can be weakened if we have other extra conditions holding.

# Chapter 4

# Preliminaries

## 4.1 Basic definitions

A *tree poset* is a partial order with a maximum element, called the root, such that every element that is not the root has exactly one upper cover. The minimal elements of the partial order are called the *leaves*. The Hasse diagram of a tree poset is a tree in the graph-theoretic sense; here we use the word *tree* as a synonym for tree poset. For $p \geq 2$, a *p-ary tree* is a tree where every element has at most $p$ lower covers. We will use *binary* and *ternary* as synonyms for 2-ary and 3-ary, respectively. A *complete p-ary tree* is a $p$-ary tree such that all maximal chains are of equal length and every element that is not a leaf has exactly $p$ lower covers. The height of a complete $p$-ary tree fully determines the partial order, for example the complete $p$-ary tree of height $n$, denoted by $T_p^n$, has $(p^n - 1)/(p - 1)$ elements and $p^{n-1}$ leaves.

For $T$ a tree, we write $1_T$ for the root of $T$, and write $1_n$ for the root of $T_p^n$. An embedding $\phi$ of a tree $T$ into the complete tree $T_p^n$ is a map from $T$ to $T_p^n$ such that $\phi(x) > \phi(y)$ in $T_p^n$ if and only if $x > y$ in $T$. Define $A_T^{(p)}(n)$ to be the number of embeddings $\phi$ of $T$ into $T_p^n$ with $\phi(1_T) = \phi(1_n)$ and define $C_T^{(p)}(n)$ to be the total number of embeddings of $T$ into $T_p^n$.

## 4.2   Background

This work is motivated by results from previous papers by Kubicki, Lehel and Morayne. In [22], Morayne looked at a partial order analogue to the secretary problem. For a detailed history and discussion of the secretary problem see [13]. As explained there, the classical secretary problem is to find the optimal strategy given the following set-up.

1. There is one secretarial position available.

2. The number $N$ of applicants is known.

3. The applicants are interviewed sequentially in random order, each order being equally likely.

4. It is assumed there is a ranking of the applicants from best to worst without ties. The decision to accept or reject an applicant must be based only on the relative ranks of those applicants interviewed so far.

5. An applicant once rejected cannot later be recalled.

6. Your payoff is 1 if you choose the best of the $N$ applicants and 0 otherwise.

It turns out that the optimal strategy is to reject the first $M = M(N)$ applicants and then accept the next applicant who is the best out of those already interviewed. There is an explicit expression for $M$ in terms of $N$, and $M/N \to 1/e$ as $N \to \infty$. Furthermore, the probability of success by following this strategy tends to $1/e$ as $N \to \infty$. See, for example, [15] for both exact and asymptotic results.

The above conditions can all be modified to give different variants of the problem, for example where the interviewer has $k$ offers in which to get the best secretary, or where it is possible to recall a rejected candidate (with some cost). Morayne considers the situation where condition 4 is modified so that the applicants are

ordered as a complete binary tree, and at each stage the interviewer knows the partial order formed by the applicants interviewed so far. There is still a best applicant and the problem is to find the optimal strategy that maximises the probability of choosing the best applicant. The number of applicants $N$ is equal to $2^n - 1$ for some $n$, the height of the complete binary tree $T_2^n$.

In [22], Morayne proves that the following strategy is optimal. If the partial order of the interviewed candidates is not a chain, and the current applicant is best-so-far then we accept him. If the partial order is a chain, then we only accept a best-so-far applicant if the height of the chain is greater than $n/2$, half the height of $T_2^n$. In other words, the strategy is to take the first applicant that is best-so-far, with the caveat that if the first $k$ applicants are totally ordered (for $k < n/2$), then we should not take the best-so-far. This caveat for the chain case is necessary, as can be seen in the following example. Suppose that $n$ is large so that the number of applicants $N = 2^n - 1$ is large in comparison to $n$, and suppose that after three interviews the partial order of the applicants is a chain, with the third applicant the best-so-far. It is highly unlikely that the third applicant is the best; this probability is of the order $n^{-1}$ and tends to 0 as $n$ tends to infinity. In contrast, if after three interviews we have that the third applicant is best-so-far, and the first two are incomparable, then the probability that the third applicant is the best is greater than $1/2$, for any $n \geq 2$.

In this set-up, where the applicants are ordered as a complete binary tree, let us look at the probability that the current applicant is the best applicant. As is the case for the classical secretary problem, if the current applicant is not the best-so-far then they cannot be the best applicant. So, consider the case when the current applicant *is* the best-so-far. Therefore, the partial order seen by the interviewer is a tree with the maximum element of the tree being the current applicant. Denote this tree by $T$. The fact that the interviewer knows the ordering of the interviewed applicants after *each* interview means we can consider $T$ to be a labelled tree. Sup-

pose we have interviewed $k$ applicants, so that $T$ is a $k$-element tree labelled with the numbers 1 to $k$, and the root is labelled with $k$. To calculate the probability that the current applicant is the best, given the tree $T$, we need to count the number of orderings of the elements of $T_2^n$ so that the first $k$ form an isomorphic copy of $T$, and count the proportion of those that have the $k$-th element as the root of $T_2^n$. Put another way, we count the number of labellings of $T_2^n$ (with the numbers $1 \ldots, N$) for which the elements labelled $1, \ldots, k$ form an isomorphic copy of $T$, and count the proportion of those which have the root of $T_2^n$ labelled $k$. But this is exactly the proportion of embeddings of $T$ into $T_2^n$ that map the root of $T$ to the root of $T_2^n$. Therefore, the probability that the current applicant is the best applicant, given that the interviewed applicants form the tree $T$, is the ratio $A_T^{(2)}(n)/C_T^{(2)}(n)$.

Morayne shows that this probability is greater than $1/2$ when either $T$ is not a chain, or $T$ is a chain with $k > n/2$. This means that any strategy which dictates that we should continue interviewing applicants, given such a $T$, is a worse strategy. Morayne also shows that to stop interviewing if $T$ is a chain with $k \leq n/2$ is not an optimal strategy, which essentially shows the strategy given above is optimal.

Important to Morayne's proof is the fact that, for $T$ a chain, $A_T^{(2)}(n)/C_T^{(2)}(n)$ is increasing in the height of the chain. This leads naturally to the question: Is the ratio $A_T^{(2)}(n)/C_T^{(2)}(n)$ "increasing in $T$", for other trees $T$? In other words, if $T_1, T_2$ are trees with $T_1$ a subtree of $T_2$ does the inequality

$$\frac{A_{T_1}^{(2)}(n)}{C_{T_1}^{(2)}(n)} \leq \frac{A_{T_2}^{(2)}(n)}{C_{T_2}^{(2)}(n)} \tag{4.1}$$

hold? The intuition is that for the larger tree, we have more information about the current applicant—he is better than more applicants—so the probability that he is the best applicant should be greater. As mentioned above, in [22] Morayne proves the inequality when $T_1, T_2$ are chains. In [17], Kubicki, Lehel and Morayne prove the inequality for binary trees $T_1, T_2$, as below.

**Theorem 4.1** (Kubicki, Lehel and Morayne). *For any $n$ and any binary trees $T_1, T_2$*

*with $T_1$ a subposet of $T_2$ we have*

$$\frac{A_{T_1}^{(2)}(n)}{C_{T_1}^{(2)}(n)} \leq \frac{A_{T_2}^{(2)}(n)}{C_{T_2}^{(2)}(n)}.$$

They prove the result using lemmas which can be stated as follows.

**Lemma 4.2.** *For any binary tree $T$, and any $x$ in $T$, we write $C_T^{(2)}(n; x \to k)$ for the number of embeddings of $T$ into $T_2^n$ that map $x$ to an element in level $k$ of $T_2^n$, and write $A_T^{(2)}(n; x \to k)$ for the number of those embeddings that also map the root of $T$ to the root of $T_2^n$. We have*

$$\frac{A_T^{(2)}(n; x \to k)}{C_T^{(2)}(n; x \to k)} \leq \frac{A_T^{(2)}(n; x \to k+1)}{C_T^{(2)}(n; x \to k+1)}. \tag{4.2}$$

**Lemma 4.3.** *For any binary tree $T$, in both cases*

*(a) $x$ a leaf of $T$,*

*(b) $x$ the only lower cover of an element in $T$,*

*we have*

$$\frac{A_T^{(2)}(k)}{A_S^{(2)}(k)} \leq \frac{A_T^{(2)}(k+1)}{A_S^{(2)}(k+1)}, \tag{4.3}$$

*where $S = T \setminus \{x\}$.*

Both lemmas are proved by "brute force"; Kubicki, Lehel and Morayne calculate expressions for each term in the inequalities (4.2) and (4.3) in terms of sums of products of similar expressions for smaller trees, and then apply induction to get the result. They combine the two lemmas and results about log-concavity of sequences to prove Theorem 4.1.

Informally, Lemma 4.2 states that, for any element $x$ in $T$, if an embedding maps $x$ to a higher level (nearer to the root) of $T_2^n$ then it is more likely that the embedding also maps the root of $T$ to a higher level, e.g., to the root. Lemma 4.3 states that, in both of the two cases stated, there are proportionally more embeddings of the

larger tree $T$ than of $S = T \setminus \{x\}$ when the root is mapped to a higher level of $T_2^n$. Both the lemmas give the impression of being correlation inequalities (on some unspecified probability space). Indeed, we show in chapter 8 that the lemmas are examples of correlation inequalities on certain lattices. In this way we can generalise Theorem 4.1 to embeddings, and other mappings, into more general complete trees. We essentially follow the proof method of Kubicki, Lehel and Morayne in [17], but using the power of the FKG-inequality we can simplify (and therefore more easily generalise) Lemmas 4.2 and 4.3.

In [17], Kubicki, Lehel and Morayne conjectured that the inequality (4.1) also holds for arbitrary trees, as follows.

**Conjecture 4.4** (Kubicki, Lehel and Morayne). *For any $n$ and any trees $T_1$, $T_2$ with $T_1$ a subtree of $T_2$ we have*

$$\frac{A_{T_1}^{(2)}(n)}{C_{T_1}^{(2)}(n)} \leq \frac{A_{T_2}^{(2)}(n)}{C_{T_2}^{(2)}(n)}.$$

In [19], Kubicki, Lehel and Morayne show that Conjecture 4.4 is true for stars rooted at their centre. Intuitively the conjecture seems highly plausible, especially given the interpretation of the inequality in terms of the secretary problem. Surprisingly then, the conjecture is in fact false. Indeed, we will show that the conjecture is false even when we restrict $T_1, T_2$ to being ternary. As mentioned in the introduction, we study the asymptotics of $A_T^{(2)}(n)/C_T^{(2)}(n)$ which helps direct our search for counterexamples to Conjecture 4.4 for arbitrarily large $n$.

We begin with some recurrence relations for $A_T^{(2)}(n)$ and $C_T^{(2)}(n)$. For ease of notation we write $T^n$ for the complete binary tree $T_2^n$ and write $A_T(n), C_T(n)$ for the numbers $A_T^{(2)}(n), C_T^{(2)}(n)$.

# Chapter 5

# The expressions $A_T(n)$ and $C_T(n)$

## 5.1 Recurrence relations for $A_T(n)$ and $C_T(n)$

We can use the regular structure of $T^n$ to find recurrence relations for $A_T(n)$ and $C_T(n)$. Let $t_1, t_2$ be the 2 lower covers of $1_n$ in $T^n$. Write $(T^n)_1$ for the set of all elements that are lower than or equal to $t_1$ in $T^n$, and similarly for $(T^n)_2$. So, $(T^n)_1$ and $(T^n)_2$ are both copies of $T^{n-1}$. For any embedding of a tree $T$ into $T^n$ the root $1_T$ of $T$ is either mapped to $1_n$, or mapped into $(T^n)_1$ or $(T^n)_2$. Counting these embeddings of $T$ into $T^n$ gives

$$C_T(n) - 2C_T(n-1) = A_T(n). \tag{5.1}$$

So, once we have calculated $A_T(n)$ we can solve a simple linear recurrence to find $C_T(n)$.

We now show that $A_T(n)$ also satisfies a linear recurrence relation. For any $x \in T$ we write $D[x]$ for the set of all elements in $T$ that are lower than or equal to $x$ in $T$. Let $T$ be a tree and suppose the root $1_T$ has $r$ lower covers $x_1, \ldots, x_r$. For any subset $L \subseteq [r]$ write $T_L$ for the tree formed by removing the subtrees $D[x_j]$ for all $j \in L^c$. (Here, $L^c = [r] \setminus L$.) Notice that $T_{\{j\}} \setminus \{1_T\} = D[x_j]$, $T_{[r]} = T$ and $T_\emptyset = \{1_T\}$.

We will count the embeddings of $T$ into $T^n$ by considering the possible places to map the elements $x_1, \ldots, x_r$. In particular we are interested in the partition of $\{x_1, \ldots, x_r\}$ defined by which of the two subtrees $(T^n)_1, (T^n)_2$ an element $x_i$ is mapped to.

Write $A^-_{T_L}(n)$ for the number of embeddings of $T_L$ into $T^n$ that map the root $1_T$ of $T_L$ to $1_n$ and map $x_j$ into $(T^n)_1$, for each $j \in L$. By the symmetry of $T^n$ this is the same as the number of embeddings of $T_L$ into $T^n$ that map $1_T$ to $1_n$ and map $x_j$ into $(T^n)_2$, for each $j \in L$.

For a fixed set $L \subseteq [r]$ we can count the number of embeddings $\phi$ of $T$ into $T^n$ with $\phi(x_i)$ in $(T^n)_1$ for all $i \in L$, and $\phi(x_i)$ in $(T^n)_2$ for all $i \in L^c$. Since the two trees $(T^n)_1$ and $(T^n)_2$ are below incomparable elements $t_1$ and $t_2$, we have that the number of such embeddings that also map $1_T$ to $1_n$ is exactly the product $A^-_{T_L}(n)A^-_{T_{L^c}}(n)$. So,

$$A_T(n) = \sum_{L \subseteq [r]} A^-_{T_L}(n)A^-_{T_{L^c}}(n). \tag{5.2}$$

For $L = \emptyset$, we have $T_\emptyset = \{1_T\}$ and $A^-_{T_L}(n)$ is equal to 1. For $L$ a singleton, $A^-_{T_L}(n)$ is the number of embeddings of $T_L \setminus \{1_T\} = D[x_j]$ into $(T^n)_1$, which itself is a copy of $T^{n-1}$. So $A^-_{T_L}(n) = C_{D[x_j]}(n-1)$. Finally, for $|L| \geq 2$, $A^-_{T_L}(n)$ is the number of embeddings that map $1_T$ to $1_n$ and map $x_j$ to an element of $(T^n)_1$ for all $j \in L$. Since $|L| \geq 2$ any such embedding $\phi$ cannot map any of the $x_j$ to $t_1$. So, for each embedding $\phi$ we can construct a new embedding $\psi$ of $T_L$ into $T^n$ by defining $\psi(1_T) = t_1$ and $\psi(x) = \phi(x)$ for all $x \in T_L \setminus \{1_T\}$. Now, $\psi$ is an embedding into $(T^n)_1$ which maps $1_T$ to $t_1$, the root of $(T^n)_1$. Since $(T^n)_1$ is a copy of $T^{n-1}$ the number of these embeddings $\psi$ is $A_{T_L}(n-1)$. Since each $\phi$ corresponds uniquely to a $\psi$, and vice-versa, we must have $A^-_{T_L}(n) = A_{T_L}(n-1)$. To summarise,

$$A^-_{T_L}(n) = \begin{cases} 1 & L = \emptyset \\ C_{D[x_j]}(n-1) & L = \{j\} \\ A_{T_L}(n-1) & \text{otherwise.} \end{cases} \tag{5.3}$$

It will also be useful to have another expression for $A_{T_L}^-(n)$ when $L = \{j\}$. We have that $A_{T_L}^-(n)$ is the number of embeddings of $T_L$ into $T^n$ that map $1_T$ to $1_n$ and map $x_j$ to an element in $(T^n)_1$. By symmetry of $T^n$ it is also the number of embeddings of $T_L$ into $T^n$ that map $1_T$ to $1_n$ and map $x_j$ to an element in $(T^n)_2$. Since, every embedding of $T_L$ into $T^n$ that maps $1_T$ to $1_n$ must map $x_j$ to an element in either $(T^n)_1$ or $(T^n)_2$ we have $2A_{T_L}^-(n) = A_{T_L}(n)$ or

$$A_{T_L}^-(n) = \frac{A_{T_L}(n)}{2} \tag{5.4}$$

for $L = \{j\}$.

We can use equations (5.1)–(5.4) to find $A_T(n)$ and $C_T(n)$ inductively. For $T$ a tree, the number of leaves of $T$ is denoted by $l(T)$.

**Theorem 5.1.** *For any tree $T$, the number of embeddings of $T$ into $T^n$ is of the form*

$$C_T(n) = \sum_{j=0}^{l(T)} g_j(n) 2^{jn},$$

*where each $g_j$ is a polynomial.*

*For $T$ the 1-element tree, the number of these embeddings that map the root of $T$ to $1_n$, $A_T(n)$, is equal to 1. Otherwise, for $T$ with $|T| > 1$, the number is of the form*

$$A_T(n) = \sum_{j=0}^{l(T)} q_j(n) 2^{jn},$$

*where each $q_j$ is a polynomial.*

The following lemma on recurrence relations will be useful.

**Lemma 5.2.** *Suppose $l$ is some fixed positive integer. Then the solution to the equation*

$$y_n - 2y_{n-1} = \sum_{j=0}^{l} f_j(n) 2^{jn}, \quad y_1 = 0, \tag{5.5}$$

*where each $f_j$ is a polynomial, is*

$$y_n = \sum_{j=0}^{l} g_j(n) 2^{jn}$$

*where each $g_j$ is a polynomial. Furthermore, for $j \neq 1$, the polynomial $g_j$ is the unique polynomial satisfying the identity*

$$g_j(n) - 2^{1-j}g_j(n-1) = f_j(n),$$

*and $g_1$ satisfies the identity*

$$g_1(n) - g_1(n-1) = f_1(n),$$

*where the constant term of $g_1$ is given by*

$$\sum_{j=0}^{l} g_j(1)2^j = 0$$

**Proof.** By linearity, it is enough to find the complementary solution, and the particular solutions to $y_n - 2y_{n-1} = f_j(n)2^{jn}$ for each $j$. The complementary solution is the solution to $y_n - 2y_{n-1} = 0$, which is just $y_n = K2^n$, for some constant $K$. For a fixed $j$, a particular solution to $y_n - 2y_{n-1} = f_j(n)2^{jn}$ is of the form $y_n = h_j(n)2^{jn}$, for some polynomial $h_j$, by an elementary result on recurrence relations. For $j \neq 1$, the polynomial $h_j$ has the same degree as $f_j$, and $h_1$ has degree one larger than $f_1$. (Also, assume $h_1$ has no constant term; this is covered by the complementary solution.) The general solution is $y_n = \sum_{j=0}^{l} g_j(n)2^{jn}$, where $g_j(n) = h_j(n)$ for all $j \neq 1$ and $g_1(n) = h_1(n) + K$. So, the solution to (5.5) will be of the required form.

Moreover, if $y_n = h_j(n)2^{jn}$ is the particular solution to $y_n - 2y_{n-1} = f_j(n)2^{jn}$, then we have $h_j(n) - 2^{1-j}h_j(n-1) = f_j(n)$. So, for all $j$, the polynomial $g_j$ satisfies

$$g_j(n) - 2^{1-j}g_j(n-1) = f_j(n).$$

For $j \neq 1$, since $g_j$ and $f_j$ are polynomials of the same degree, this equation uniquely determines $g_j$. Since the equation $h_1(n) - h_1(n-1) = f_1(n)$ uniquely determines $h_1$ (as we assumed that $h_1$ has no constant term), the polynomial $g_1$ is determined except for the constant term $K$. We fix $K$ with the initial condition of (5.5), $y_1 = 0$. This gives the equation

$$\sum_{j=0}^{l} g_j(1)2^j = 0$$

as claimed. □

**Proof of Theorem 5.1.** We include the case of $T$ being a 1-element set for completeness. In this case, we see immediately that there are $2^n - 1$ embeddings of $T$ into $T^n$, which is exactly the number of elements in $T^n$. Also, only one of these embeddings maps the root of $T$ to $1_n$. So, $A_T(n) = 1$ as claimed, and $C_T(n) = 2^n - 1$ is of the required form.

For $|T| \geq 2$, we simultaneously prove that $A_T(n)$ and $C_T(n)$ are of the required form by induction on the size of $T$. We shall make use of Lemma 5.2 to solve recurrence relations for $A_T(n)$ and $C_T(n)$. We use induction to show that the recurrence is of the form of equation (5.5), and since we will only be considering trees with $|T| \geq 2$ we have the initial conditions $A_T(1) = 0, C_T(1) = 0$ as in (5.5).

For $|T| = 2$ the only tree is the 2-element chain, which has one leaf. Label the root $1_T$ and the leaf $x_1$. Since $1_T$ has only one lower cover, $r = 1$ in equation (5.2) and the subtrees of interest are $T_{\{1\}} = T$ and $T_{\emptyset} = \{1_T\}$. Using equations (5.2) and (5.3) we have

$$A_T(n) = A^-_{T_\emptyset}(n)A^-_{T_{\{1\}}}(n) + A^-_{T_{\{1\}}}(n)A^-_{T_\emptyset}(n) = 2C_{\{x_1\}}(n-1)$$

But we have shown earlier that $C_{\{x_1\}}(n) = 2^n - 1$. Therefore $A_T(n) = 2^n - 2$ which is of the required form (where $l(T) = 1$, $q_0(n) = -2$ and $q_1(n) = 1$).

In fact, we can see immediately that $A_T(n) = 2^n - 2$, since this is exactly the number of places to embed $x_1$ in $T^n$ (anywhere except at $1_n$, where $x$ is embedded). Using (5.1) and Lemma 5.2 we have that $C_T(n) = (n-2)2^n + 2$ which is of the required form ($g_1(n) = n - 2$ and $g_0(n) = 2$).

Suppose the result is true for all $T$ with $|T| < k$ and let $T$ be any tree with $|T| = k$. There are two cases to consider, depending on whether the root of $T$ has exactly one lower cover. If the root has exactly one lower cover, $x_1$, equation (5.2)

reduces, in a similar way to the base case, to

$$A_T(n) = 2C_{D[x_1]}(n-1).$$

Applying the inductive hypothesis to $D[x_1]$, a tree with $l(D[x_1]) = l(T)$ leaves, we have that

$$C_{D[x_1]}(n) = \sum_{j=0}^{l(T)} g_j(n)2^{jn}$$

where $g_j$ are polynomials. Therefore,

$$A_T(n) = 2\sum_{j=0}^{l(T)} g_j(n-1)2^{j(n-1)} = \sum_{j=0}^{l(T)} q_j(n)2^{jn}$$

where $q_j$ are polynomials.

If the root of $T$ has $r > 1$ lower covers $x_1, \ldots, x_r$ then we can write equation (5.2) as

$$A_T(n) = A_{T_\emptyset}^-(n)A_{T_{[r]}}^-(n) + A_{T_{[r]}}^-(n)A_{T_\emptyset}^-(n) + \sum_{\substack{L \subseteq [r] \\ L \neq \emptyset, [r]}} A_{T_L}^-(n)A_{T_{L^c}}^-(n)$$

which can be rearranged to

$$A_T(n) - 2A_T(n-1) = \sum_{\substack{L \subseteq [r] \\ L \neq \emptyset, [r]}} A_{T_L}^-(n)A_{T_{L^c}}^-(n). \tag{5.6}$$

We use equations (5.3) and (5.4) in order to apply the inductive hypothesis. Terms in the sum where $L$ is not a singleton or complement of a singleton are of the form $A_{T_L}(n-1)A_{T_{L^c}}(n-1)$. Terms where $L$ is a singleton but $L^c$ is not, are of the form $A_{T_L}(n)A_{T_{L^c}}(n-1)/2$; terms where $L$ is not a singleton but $L^c$ is, are of the form $A_{T_L}(n-1)A_{T_{L^c}}(n)/2$ and terms where both $L$ and $L^c$ are singletons (this will only be for $r = 2$) are of the form $A_{T_L}(n)A_{T_{L^c}}(n)/4$.

By our inductive hypothesis we have $A_{T_L}(n) = \sum_{j=0}^{l(T_L)} q_j(n)2^{jn}$ for polynomials $q_j$. This means that the right hand side of equation (5.6) is of the form $\sum_{j=0}^{l(T)} h_j(n)2^{jn}$ for polynomials $h_j$. That is, $A_T(n)$ satisfies a recurrence relation and applying Lemma 5.2 gives the result for $A_T(n)$. Finally, we use (5.1) and Lemma 5.2 which gives the result for $C_T(n)$. $\qquad\square$

## 5.2 Counterexamples to a conjecture of Kubicki, Lehel and Morayne

Note that the proof of Theorem 5.1 actually shows how to find the polynomials $q_j$ and $g_j$ in the expressions for $A_T(n)$ and $C_T(n)$. However, for a particular tree $T$, in order to calculate $A_T(n)$ and $C_T(n)$ we need to calculate $A_{T_L}(n)$ for all subtrees $T_L$. For small trees the calculations are still relatively simple. We use the algorithm given in the proof of Theorem 5.1 to find explicit expressions for the two trees $T_1, T_2$ in Figure 5.1.
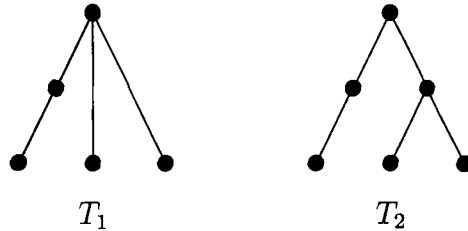


Figure 5.1: Counterexample to Conjecture 4.4

To find these expressions we need to also calculate $A_S$ and $C_S$ for subtrees $S$ of $T_1$ and $T_2$. Define the subtrees $S_1 = \;$, $S_2 = \;$, $S_3 = \;$, $S_4 = \;$, $S_5 = \;$. In order to find $A_{T_1}$ we need to calculate $A_{S_1}, A_{S_2}, A_{S_3}, A_{S_4}$, and to find $A_{S_1}$ we need to calculate $C_{S_2}$. For $A_{T_2}$ we also need to calculate $A_{S_5}$ and to find this we need to calculate $C_{S_4}$. Table 5.1 lists the expressions $A_S(n), C_S(n)$ needed.

Solving the recurrence relations for $T_1$ and $T_2$, using the expressions in Table 5.1, gives

$$A_{T_1}(n) = (n - 14/3)8^n + (-3n^2 + 24n - 34)4^n + (n^3/3 - 8n^2 + 65n/3 + 44/3)2^n + 24$$

$$A_{T_2}(n) = (2n/3 - 20/9)8^n + (-n^3 + 8n^2 - 30n + 58)4^n$$
$$+ (-2n^3/3 + 2n^2 - 40n/3 - 430/9)2^n - 8$$

| $S$ | $A_S(n)$ | $C_S(n)$ |
|---|---|---|
| (path) | $2^n - 2$ | $(n-2)2^n + 2$ |
| (path) | $(n-3)2^n + 4$ | $(n^2/2 - 5n/2 + 4)2^n - 4$ |
| (tree) | $4^n + (-2n+1)2^n - 2$ | $2.4^n + (-n^2 - 4)2^n + 2$ |
| (tree) | $(n-4)4^n + (-n^2/2 + 9n/2)2^n + 4$ | Not needed |
| (tree) | $4^n + (-n^2 + 2n - 5)2^n + 4$ | Not needed |

Table 5.1: $A_S(n), C_S(n)$ for small trees $S$

and, using (5.1), we have

$$C_{T_1}(n) = (4n/3 - 20/3)8^n + (-6n^2 + 60n - 134)4^n$$

$$+ (n^4/12 - 5n^3/2 + 83n^2/12 + 145n/6 + 494/3)2^n - 24$$

$$C_{T_2}(n) = (8n/9 - 88/27)8^n + (-2n^3 + 22n^2 - 110n + 250)4^n$$

$$+ (-n^4/6 + n^3/3 - 35n^2/6 - 487n/9 - 6878/27)2^n + 8$$

So, $A_{T_1}(4)/C_{T_1}(4) = 99/101 > 67/69 = A_{T_2}(4)/C_{T_2}(4)$, a counterexample to the conjecture of Kubicki, Lehel and Morayne. We also have

$$\frac{A_{T_1}(5)}{C_{T_1}(5)} = \frac{2635}{2837} > \frac{1783}{1921} = \frac{A_{T_2}(5)}{C_{T_2}(5)}$$

but

$$\frac{A_{T_1}(6)}{C_{T_1}(6)} = \frac{44147}{49821} < \frac{31055}{34897} = \frac{A_{T_2}(6)}{C_{T_2}(6)}.$$

So, for $n = 4, 5$ these trees give a counterexample, but not for $n = 6$. In fact, for $n = 6, \ldots, 11$ the conjectured inequality holds, but for larger $n$ it does not. Asymptotically, we have

$$\frac{A_{T_1}(n)}{C_{T_1}(n)} = \frac{(n - 14/3)8^n + O(n^2 4^n)}{(4n/3 - 20/3)8^n + O(n^2 4^n)} = \frac{3}{4} + \frac{1}{4}n^{-1} + \frac{5}{4}n^{-2} + o(n^{-2})$$

and

$$\frac{A_{T_2}(n)}{C_{T_2}(n)} = \frac{(2n/3 - 20/9)8^n + O(n^3 4^n)}{(8n/9 - 88/27)8^n + O(n^3 4^n)} = \frac{3}{4} + \frac{1}{4}n^{-1} + \frac{11}{12}n^{-2} + o(n^{-2}),$$

so $A_{T_1}/C_{T_1}$ is asymptotically larger than $A_{T_2}/C_{T_2}$.

This asymptotic difference is very subtle. Here, the ratios $A_{T_1}/C_{T_1}$, $A_{T_2}/C_{T_2}$ differ only in the $n^{-2}$ terms and terms of lower order. We will show, in Section 6.3, that for any $T_1 \subseteq T_2$ which have $A_{T_1}/C_{T_1}$ asymptotically larger than $A_{T_2}/C_{T_2}$ the ratios differ only in the $n^{-2}$ terms and terms of lower order.

For small values of $n$ there are two competing factors which determine whether the conjectured inequality holds. Since $A_T$ and $C_T$ are related by (5.1), we have $A_T(n)/C_T(n) = 1 - 2C_T(n-1)/C_T(n)$. So, the conjectured inequality is equivalent to

$$\frac{C_{T_2}(n-1)}{C_{T_1}(n-1)} \leq \frac{C_{T_2}(n)}{C_{T_1}(n)}.$$

We can think of the ratio $C_{T_2}(n)/C_{T_1}(n)$ as the expected number of embeddings of $T_2$ into $T^n$ that are an extension of a randomly chosen embedding of $T_1$ into $T^n$. So, for $n = 3$, each embedding of $T_1$ into $T^3$ can only be extended one way (there is only one place in $T^3$ to which we can map the extra element of $T_2$), therefore $C_{T_2}(3)/C_{T_1}(3) = 1$. For larger values of $n$, some embeddings of $T_1$ into $T^n$ have no extensions to an embedding of $T_2$ into $T^n$, others will have many extensions to an embedding of $T_2$ into $T^n$. In this example, as $n$ increases there will tend to be a larger fraction of embeddings of $T_1$ into $T^n$ with no extension to an embedding of $T_2$ into $T^n$. However, those embeddings of $T_1$ into $T^n$ that do have extensions to embeddings of $T_2$ into $T^n$ will tend to have more of them, as $n$ increases. These two competing effects determine whether the ratio $C_{T_2}(n)/C_{T_1}(n)$ will increase or decrease for an increase in $n$. In this example the two effects are quite equally balanced, making it difficult to see intuitively why the inequality holds for some values of $n$ and fails for others.

The following example better illustrates the failure of the conjectured inequality,
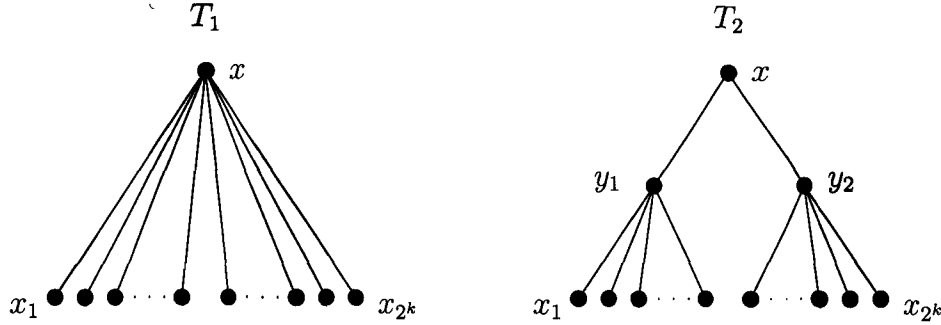
Figure 5.2: Counterexample to Conjecture 4.4

as in this example one effect dominates the other. Let $T_1$ and $T_2$ be as shown in Figure 5.2, where $k$ is some fixed integer. As we have explained, the conjecture claims that $C_{T_2}(n)/C_{T_1}(n)$ is increasing in $n$. However, we show that for these trees, the ratio is considerably larger for small $n$ than it is for large $n$, since for small $n$ there is a higher proportion of embeddings of $T_1$ that can be extended to an embedding of $T_2$.

For any $n$ with $n \geq k + 1$, an embedding of $T_2$ into $T^n$ must map all the leaves $x_1, \ldots, x_{2^{k-1}}$ into the same half of $T^n$, and it must map all the leaves $x_{2^{k-1}+1}, \ldots, x_{2^k}$ into the same half of $T^n$. This is a restriction imposed by the elements $y_1$ and $y_2$. Embeddings of $T_1$ into $T^n$ do not have this restriction, and any embedding of $T_1$ into $T^n$, which does not partition the leaves in the same way cannot be extended to an embedding of $T_2$.

Now, for $n = k+1$, the tree $T^{k+1}$ has $2^k$ leaves, so all embeddings of $T_1$ into $T^{k+1}$ map the leaves of $T_1$ to the leaves of $T^{k+1}$. Therefore, we know that half the leaves of $T_1$ are mapped into one half of $T^{k+1}$ and the other half into the other half of $T^{k+1}$. So whether the embedding extends to an embedding of $T_2$ depends only on which particular set of $2^{k-1}$ leaves are mapped into one of the halves of $T^{k+1}$. Since there are $\binom{2^k}{2^{k-1}}$ subsets of size $2^{k-1}$, and two of these yield an extendible embedding (when we choose $\{x_1, \ldots, x_{2^{k-1}}\}$ or $\{x_{2^{k-1}+1}, \ldots, x_{2^k}\}$), each with one possible extension, the ratio $C_{T_2}(k+1)/C_{T_1}(k+1)$ is equal to $2/\binom{2^k}{2^{k-1}}$.

For $n \gg k+1$, most mappings from $T_1$ into $T^n$ are embeddings, but only those which partition the leaves as described above can be extended. Moreover, most of the embeddings that can be extended map the leaves $x_1, \ldots, x_{2^{k-1}}$ into one half of $T^n$, and the leaves $x_{2^{k-1}+1}, \ldots, x_{2^k}$ into the other half of $T^n$ (rather than the same half) and most of these extendible embeddings have only one possible extension. So of the total number of embeddings of $T_1$ into $T^n$ the fraction that are extendible is roughly $2^{-2^k}$ and most extendible embeddings have just one possible extension. Therefore, $C_{T_2}(n)/C_{T_1}(n)$ is roughly $1/2^{2^k}$, which is considerably smaller than $C_{T_2}(k+1)/C_{T_1}(k+1) = 2/\binom{2^k}{2^{k-1}}$.

# Chapter 6

# Asymptotic behaviour of $A_T(n)$ and $C_T(n)$

In this chapter we study the asymptotic behaviour of $A_T(n)$ and $C_T(n)$ in order to provide counterexamples to Conjecture 4.4, for arbitrarily large $n$. This tells us that we cannot hope for a version of the conjecture that holds "for sufficiently large $n$". The calculations are similar in style to those in the previous chapter, but here we need to be more exact, as we will need to calculate the leading terms of $A_T(n)$ and $C_T(n)$. Also, using these expressions, we are able to describe a "typical" embedding of $T$ into $T^n$ (for large $n$).

## 6.1   Leading terms of $A_T(n)$

We have shown that $A_T(n) = \sum_{j=0}^{l} q_j(n) 2^{jn}$, where each $q_j$ is a polynomial. We wish to examine the asymptotic behaviour of $A_T(n)$ and so we need to calculate the leading terms of the dominant polynomial $q_l(n)$. Throughout this chapter we use the symbol $\sim$ to mean "asymptotically equivalent to"; we write $f(n) \sim g(n)$ if $f(n)/g(n)$ tends to 1 as $n$ tends to infinity. We shall make use of the following lemma which gives the solutions to some particular recurrence relations.

**Lemma 6.1.** *The recurrence relation*

$$y_n - 2y_{n-1} = \sum_{j=0}^{l} f_j(n) 2^{jn},$$

*where each $f_j$ is a polynomial, and the leading term of $f_l(n)$ is $\alpha n^d$, has solution*

$$y_n \sim \begin{cases} \dfrac{\alpha}{d+1} n^{d+1} 2^n & \text{if } l = 1 \\[4mm] \dfrac{2^{l-1}}{2^{l-1}-1} \alpha n^d 2^{ln} & \text{if } l \geq 2. \end{cases} \tag{6.1}$$

*Furthermore, if $d > 0$ and the leading two terms of $f_l(n)$ are $\alpha n^d + \beta n^{d-1}$, then the solution is*

$$y_n \sim \begin{cases} \left( \dfrac{\alpha}{d+1} n^{d+1} + \left( \dfrac{\beta}{d} + \dfrac{\alpha}{2} \right) n^d \right) 2^n & \text{if } l = 1 \\[4mm] \dfrac{2^{l-1}}{2^{l-1}-1} \left( \alpha n^d + \left( \beta - \dfrac{d\alpha}{2^{l-1}-1} \right) n^{d-1} \right) 2^{ln} & \text{if } l \geq 2. \end{cases} \tag{6.2}$$

**Proof.** We have, for example from Lemma 5.2, that the solution to the recurrence relation is

$$y_n = \sum_{j=0}^{l} g_j(n) 2^{jn},$$

where each $g_j$ is a polynomial satisfying

$$g_j(n) - 2^{1-j} g_j(n-1) = f_j(n). \tag{6.3}$$

The dominant terms of the solution come from the polynomial $g_l$. It is a simple exercise to check, using (6.3), that if the leading term of $f_l(n)$ is $\alpha n^d$, then the leading term of $g_l(n)$ is given by (6.1), and, for $d > 0$, if the leading two terms of $f_l(n)$ are $\alpha n^d + \beta n^{d-1}$, then the leading two terms of $g_l(n)$ are given by (6.2).  $\square$

**Theorem 6.2.** *The leading polynomial $q_{l(T)}(n)$ in the expression*

$$A_T(n) = \sum_{j=0}^{l(T)} q_j(n) 2^{jn}$$

has degree $d(T)$, where $d(T) = |\{x \in T : x \text{ not the root or a leaf, } D[x] \text{ is a chain}\}|$. The coefficient $\alpha_T$ of $n^{d(T)}$ satisfies the following equations.

If $T$ is the 2-element chain, then $\alpha_T = 1$. Otherwise, if the root of $T$ has $r$ lower covers, then

$$
\alpha_T = \begin{cases}
\dfrac{\alpha_{D[x_1]}}{d(T)} & T \text{ a chain, } \quad r = 1 \\[2em]
\dfrac{\alpha_{D[x_1]}}{2^{l(T)-1} - 1} & T \text{ not a chain, } \quad r = 1 \\[2em]
\dfrac{\alpha_{T_{\{1\}}} \alpha_{T_{\{2\}}} 2^{l(T)-2}}{2^{l(T)-1} - 1} & r = 2 \\[2em]
\dfrac{\sum_{j=1}^{r} \alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}} 2^{l(T_{\{j\}})-1} + \sum_{2 \le |L| \le r-2} \alpha_{T_L} \alpha_{T_{L^c}} 2^{-1}}{2^{l(T)-1} - 1} & r \ge 3
\end{cases}
\tag{6.4}
$$

Moreover, if $d(T) > 0$ the coefficient $\beta_T$ of $n^{d(T)-1}$ satisfies the following equations.

If $T$ is the 3-element chain, then $\beta_T = -3$. Otherwise, if the root of $T$ has $r$ lower covers, then

$$
\beta_T = \begin{cases}
\dfrac{\beta_{D[x_1]}}{d(T) - 1} - \dfrac{d(T)\alpha_T}{2} & T \text{ a chain, } \quad r = 1 \\[2em]
\dfrac{\beta_{D[x_1]} - d(T)\alpha_T 2^{l(T)-1}}{2^{l(T)-1} - 1} & T \text{ not a chain, } \quad r = 1 \\[2em]
\dfrac{(\alpha_{T_{\{1\}}} \beta_{T_{\{2\}}} + \alpha_{T_{\{2\}}} \beta_{T_{\{1\}}}) 2^{l(T)-2} - d(T)\alpha_T}{2^{l(T)-1} - 1} & r = 2 \\[2em]
\dfrac{\sum_{j=1}^{r} (\alpha_{T_{\{j\}}} \beta_{T_{\{j\}^c}} + \alpha_{T_{\{j\}^c}} \beta_{T_{\{j\}}} - d(T_{\{j\}^c})\alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}}) 2^{l(T_{\{j\}})-1}}{2^{l(T)-1} - 1} \\[1em]
\quad + \dfrac{\sum_{2 \le |L| \le r-2} (\alpha_{T_L} \beta_{T_{L^c}} + \alpha_{T_{L^c}} \beta_{T_L} - d(T)\alpha_{T_L} \alpha_{T_{L^c}}) 2^{-1} - d(T)\alpha_T}{2^{l(T)-1} - 1} & r \ge 3
\end{cases}
\tag{6.5}
$$

where $\beta_S = 0$ for any subtree $S \subseteq T$ with $d(S) = 0$.

**Proof.** We proceed by induction on $|T|$. We first show that the degree of $q_{l(T)}$ is

$d(T)$ and that $\alpha_T$ is as claimed. For $|T| = 2$ we have already shown that $T$ is the 2-element chain and $A_T(n) = 2^n - 2$. For this tree $d(T) = 0$, $l(T) = 1$, so $q_{l(T)}(n) = 1$ a polynomial of degree 0, with leading coefficient equal to 1. That is, $\alpha_T = 1$ as claimed.

Suppose the result is true for all $T$ with $|T| < k$ and let $T$ be any tree with $|T| = k$. As in the proof of Theorem 5.1, there are different cases to consider, depending on whether the root of $T$ has exactly one lower cover. If the root has exactly one lower cover, $x_1$, we have equation $A_T(n) = 2C_{D[x_1]}(n-1)$. But by Theorem 5.1, and our inductive hypothesis, we know that

$$A_{D[x_1]}(n) \sim \alpha_{D[x_1]} n^{d(D[x_1])} 2^{l(D[x_1])n}.$$

If $T$ is a chain, then $l(T) = l(D[x_1]) = 1$ and $d(T) = d(D[x_1]) + 1$ since the element $x_1$ contributes to $d(T)$ but not $d(D[x_1])$. So, $C_{D[x_1]}(n)$ satisfies the recurrence relation (5.1), which is of the form in Lemma 6.1 with $\alpha = \alpha_{D[x_1]}$, $d = d(D[x_1])$ and $l = l(D[x_1]) = 1$. So, by (6.1),

$$C_{D[x_1]}(n) \sim \frac{\alpha_{D[x_1]}}{d(D[x_1]) + 1} n^{d(D[x_1])+1} 2^n = \frac{\alpha_{D[x_1]}}{d(T)} n^{d(T)} 2^n.$$

So

$$A_T(n) = 2C_{D[x_1]}(n-1) \sim 2\frac{\alpha_{D[x_1]}}{d(T)}(n-1)^{d(T)} 2^{n-1} = \frac{\alpha_{D[x_1]}}{d(T)}(n-1)^{d(T)} 2^n.$$

Therefore $q_{l(T)}$ is of degree $d(T)$ and $\alpha_T = \alpha_{D[x_1]}/d(T)$, as claimed. If $T$ is not a chain, then $l(T) = l(D[x_1]) > 1$ and $d(T) = d(D[x_1])$ since the element $x_1$ does not contribute to either $d(T)$ or $d(D[x_1])$. As above, $C_{D[x_1]}(n)$ satisfies a recurrence relation of the form in Lemma 6.1 with $\alpha = \alpha_{D[x_1]}$, $d = d(T)$ and $l = l(T) > 1$. So, by (6.1),

$$C_{D[x_1]}(n) \sim \frac{2^{l(T)-1}}{2^{l(T)-1} - 1} \alpha_{D[x_1]} n^{d(T)} 2^{l(T)n}.$$

So

$$A_T(n) = 2C_{D[x_1]}(n-1) \sim 2\frac{2^{l(T)-1}}{2^{l(T)-1} - 1} \alpha_{D[x_1]}(n-1)^{d(T)} 2^{l(T)(n-1)}$$

$$= \frac{\alpha_{D[x_1]}}{2^{l(T)-1} - 1}(n-1)^{d(T)} 2^{l(T)n}.$$

Therefore $q_{l(T)}$ is of degree $d(T)$ and $\alpha_T = \alpha_{D[x_1]}/(2^{l(T)-1} - 1)$, as claimed.

If the root of $T$ has two lower covers $x_1, x_2$ then equations (5.6) and (5.4) give $A_T(n) - 2A_T(n-1) = A_{T_{\{1\}}}(n)A_{T_{\{2\}}}(n)/2$. So,

$$A_T(n) - 2A_T(n-1) \sim \alpha_{T_{\{1\}}} n^{d(T_{\{1\}})} 2^{l(T_{\{1\}})n} \alpha_{T_{\{2\}}} n^{d(T_{\{2\}})} 2^{l(T_{\{2\}})n}/2$$

$$= \alpha_{T_{\{1\}}} \alpha_{T_{\{2\}}} n^{d(T)} 2^{l(T)n}/2$$

since $d(T_{\{1\}}) + d(T_{\{2\}}) = d(T)$ and $l(T_{\{1\}}) + l(T_{\{2\}}) = l(T)$. So, $A_T(n)$ satisfies a recurrence relation of the form in Lemma 6.1 with $\alpha = \alpha_{T_{\{1\}}} \alpha_{T_{\{2\}}}/2$, $d = d(T)$, $l = l(T) > 1$. So, by (6.1),

$$A_T(n) \sim \frac{2^{l(T)-1} \alpha_{T_{\{1\}}} \alpha_{T_{\{2\}}}}{(2^{l(T)-1} - 1)2} n^{d(T)} 2^{l(T)n}$$

so $q_{l(T)}$ has degree $d(T)$ and $\alpha_T$ is as claimed.

Finally, if the root of $T$ has $r \geq 3$ lower covers $x_1, \ldots, x_r$ we can write (5.6) as

$$A_T(n) - 2A_T(n-1) = 2\sum_{j=1}^{r} \frac{1}{2} A_{T_{\{j\}}}(n) A_{T_{\{j\}^c}}(n-1)$$

$$+ \sum_{2 \leq |L| \leq r-2} A_{T_L}(n-1) A_{T_{L^c}}(n-1).$$

Terms in the first sum are of the form

$$\alpha_{T_{\{j\}}} n^{d(T_{\{j\}})} 2^{l(T_{\{j\}})n} \alpha_{T_{\{j\}^c}}(n-1)^{d(T_{\{j\}^c})} 2^{l(T_{\{j\}^c})(n-1)} \sim \frac{\alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}}}{2^{l(T_{\{j\}^c})}} n^{d(T)} 2^{l(T)n},$$

and terms in the second sum are of the form

$$\alpha_{T_L}(n-1)^{d(T_L)} 2^{l(T_L)(n-1)} \alpha_{T_{L^c}}(n-1)^{d(T_{L^c})} 2^{l(T_{L^c})(n-1)} \sim \frac{\alpha_{T_L} \alpha_{T_{L^c}}}{2^{l(T)}} n^{d(T)} 2^{l(T)n}.$$

So, $A_T(n)$ satisfies a recurrence relation of the form in Lemma 6.1 with

$$\alpha = \sum_{j=1}^{r} \frac{\alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}}}{2^{l(T_{\{j\}^c})}} + \sum_{2 \leq |L| \leq r-2} \frac{\alpha_{T_L} \alpha_{T_{L^c}}}{2^{l(T)}},$$

$d = d(T)$ and $l = l(\dot{T}) > 1$. So, by (6.1),

$$A_T(n) \sim \frac{2^{l(T)-1}}{2^{l(T)-1} - 1} \left( \sum_{j=1}^{r} \frac{\alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}}}{2^{l(T_{\{j\}^c})}} + \sum_{2 \leq |L| \leq r-2} \frac{\alpha_{T_L} \alpha_{T_{L^c}}}{2^{l(T)}} \right) n^{d(T)} 2^{l(T)n}$$

$$= \frac{\sum_{j=1}^{r} \alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}} 2^{l(T_{\{j\}})-1} + \sum_{2 \leq |L| \leq r-2} \alpha_{T_L} \alpha_{T_{L^c}} 2^{-1}}{2^{l(T)-1} - 1} n^{d(T)} 2^{l(T)n}.$$

Therefore $q_{l(T)}$ is of degree $d(T)$ and $\alpha_T$ is as claimed.

We now prove that $\beta_T$ is as claimed. For $|T| = 2$, we must have $T$ equal to the 2-element chain and so $d(T) = 0$ and there is nothing to prove. For $|T| = 3$ the only $T$ with $d(T) > 0$ is the 3-element chain. As calculated earlier (see Table 5.1), $A_T(n) = (n-3)2^n + 4$ and so $\beta_T = -3$ as claimed.

We now prove the inductive step, following the exact method used for $\alpha_T$ but we now also consider the coefficient of $n^{d(T)-1}2^{l(T)n}$ in the calculations, and use (6.2) when applying Lemma 6.1.

Suppose that $\beta_T$ is as claimed for all $T$ with $|T| < k$, that is, that $\beta_T$ satisfies (6.5) when $d(T) > 0$. Let $T$ be any tree with $|T| = k$ and $d(T) > 0$. By our inductive hypothesis we have that for all $S \subseteq T$ with $d(S) > 0$ the first two terms of $q_l(n)$ are $\alpha_S n^{d(S)} + \beta_S n^{d(S)-1}$. For some $S \subseteq T$ we may have $d(S) = 0$. For these trees set $\beta_S = 0$. Doing so means that, for all $S \subseteq T$ the first two terms of $q_l(n)$ are $\alpha_S n^{d(S)} + \beta_S n^{d(S)-1}$.

We can now consider the different cases depending on the number of lower covers of the root of $T$. If the root has exactly one lower cover, $x_1$, we have equation $A_T(n) = 2C_{D[x_1]}(n-1)$. But by Theorem 5.1, and our inductive hypothesis, we know that

$$A_{D[x_1]}(n) \sim (\alpha_{D[x_1]}n^{d(D[x_1])} + \beta_{D[x_1]}n^{d(D[x_1])-1})2^{l(D[x_1])n}.$$

If $T$ is a chain, then $l(T) = l(D[x_1]) = 1$ and $d(T) = d(D[x_1])+1$ since the element $x_1$ contributes to $d(T)$ but not $d(D[x_1])$. So, $C_{D[x_1]}(n)$ satisfies the recurrence relation (5.1), which is of the form in Lemma 6.1 with $\alpha = \alpha_{D[x_1]}$, $\beta = \beta_{D[x_1]}$, $d = d(D[x_1])$ and $l = l(D[x_1]) = 1$. So, by (6.2),

$$C_{D[x_1]}(n) \sim \left( \frac{\alpha_{D[x_1]}}{d(D[x_1]) + 1}n^{d(D[x_1])+1} + \left( \frac{\beta_{D[x_1]}}{d(D[x_1])} + \frac{\alpha_{D[x_1]}}{2} \right) n^{d(D[x_1])} \right) 2^n$$

$$= \left( \frac{\alpha_{D[x_1]}}{d(T)}n^{d(T)} + \left( \frac{\beta_{D[x_1]}}{d(T) - 1} + \frac{\alpha_{D[x_1]}}{2} \right) n^{d(T)-1} \right) 2^n.$$

Note that, since $T$ is a chain of at least four elements, we have $d(T) > 1$ so that we

are not dividing by zero. So

$$A_T(n) = 2C_{D[x_1]}(n-1)$$

$$\sim 2\left(\frac{\alpha_{D[x_1]}}{d(T)}(n-1)^{d(T)} + \left(\frac{\beta_{D[x_1]}}{d(T)-1} + \frac{\alpha_{D[x_1]}}{2}\right)(n-1)^{d(T)-1}\right)2^{n-1}$$

$$\sim \left(\frac{\alpha_{D[x_1]}}{d(T)}n^{d(T)} - \alpha_{D[x_1]}n^{d(T)-1} + \left(\frac{\beta_{D[x_1]}}{d(T)-1} + \frac{\alpha_{D[x_1]}}{2}\right)n^{d(T)-1}\right)2^n$$

$$= \left(\frac{\alpha_{D[x_1]}}{d(T)}n^{d(T)} + \left(\frac{\beta_{D[x_1]}}{d(T)-1} - \frac{\alpha_{D[x_1]}}{2}\right)n^{d(T)-1}\right)2^n.$$

Therefore, using (6.4), we have

$$\beta_T = \frac{\beta_{D[x_1]}}{d(T)-1} - \frac{\alpha_{D[x_1]}}{2} = \frac{\beta_{D[x_1]}}{d(T)-1} - \frac{d(T)\alpha_T}{2},$$

as claimed. If $T$ is not a chain, then $l(T) = l(D[x_1]) > 1$ and $d(T) = d(D[x_1])$ since the element $x_1$ does not contribute to either $d(T)$ or $d(D[x_1])$. As above, $C_{D[x_1]}(n)$ satisfies a recurrence relation of the form in Lemma 6.1 with $\alpha = \alpha_{D[x_1]}$, $\beta = \beta_{D[x_1]}$, $d = d(T)$ and $l = l(T) > 1$. So, by (6.2),

$$C_{D[x_1]}(n) \sim \frac{2^{l(T)-1}}{2^{l(T)-1}-1}\left(\alpha_{D[x_1]}n^{d(T)} + \left(\beta_{D[x_1]} - \frac{d(T)\alpha_{D[x_1]}}{2^{l(T)-1}-1}\right)n^{d(T)-1}\right)2^{l(T)n}.$$

So

$$A_T(n) = 2C_{D[x_1]}(n-1)$$

$$\sim \frac{2^{l(T)}}{2^{l(T)-1}-1}\left[\alpha_{D[x_1]}(n-1)^{d(T)} + \left(\beta_{D[x_1]} - \frac{d(T)\alpha_{D[x_1]}}{2^{l(T)-1}-1}\right)(n-1)^{d(T)-1}\right]2^{l(T)(n-1)}$$

$$\sim \frac{1}{2^{l(T)-1}-1}\left[\begin{array}{l}\alpha_{D[x_1]}n^{d(T)} - d(T)\alpha_{D[x_1]}n^{d(T)-1} \\ + \left(\beta_{D[x_1]} - \frac{d(T)\alpha_{D[x_1]}}{2^{l(T)-1}-1}\right)n^{d(T)-1}\end{array}\right]2^{l(T)n}$$

$$= \frac{1}{2^{l(T)-1}-1}\left[\alpha_{D[x_1]}n^{d(T)} + \left(\beta_{D[x_1]} - \frac{d(T)\alpha_{D[x_1]}2^{l(T)-1}}{2^{l(T)-1}-1}\right)n^{d(T)-1}\right]2^{l(T)n}.$$

Therefore, using (6.4), we have

$$\beta_T = \frac{1}{2^{l(T)-1}-1}\left(\beta_{D[x_1]} - \frac{d(T)\alpha_{D[x_1]}2^{l(T)-1}}{2^{l(T)-1}-1}\right) = \frac{\beta_{D[x_1]}}{2^{l(T)-1}-1} - \frac{d(T)\alpha_T 2^{l(T)-1}}{2^{l(T)-1}-1},$$

as claimed.

If the root of $T$ has two lower covers $x_1, x_2$ then equations (5.6) and (5.4) give

$A_T(n) - 2A_T(n-1) = A_{T_{\{1\}}}(n)A_{T_{\{2\}}}(n)/2$. So,

$$A_T(n) - 2A_T(n-1) \sim (\alpha_{T_{\{1\}}} n^{d(T_{\{1\}})} + \beta_{T_{\{1\}}} n^{d(T_{\{1\}})-1})2^{l(T_{\{1\}})n}$$

$$\times (\alpha_{T_{\{2\}}} n^{d(T_{\{2\}})} + \beta_{T_{\{2\}}} n^{d(T_{\{2\}})-1})2^{l(T_{\{2\}})n}/2$$

$$\sim \left( \frac{\alpha_{T_{\{1\}}} \alpha_{T_{\{2\}}}}{2} n^{d(T)} + \frac{\alpha_{T_{\{1\}}} \beta_{T_{\{2\}}} + \alpha_{T_{\{2\}}} \beta_{T_{\{1\}}}}{2} n^{d(T)-1} \right) 2^{l(T)n}$$

since $d(T_{\{1\}}) + d(T_{\{2\}}) = d(T)$ and $l(T_{\{1\}}) + l(T_{\{2\}}) = l(T)$. So, $A_T(n)$ satisfies a recurrence relation of the form in Lemma 6.1 with $\alpha = \alpha_{T_{\{1\}}} \alpha_{T_{\{2\}}}/2$, $\beta = (\alpha_{T_{\{1\}}} \beta_{T_{\{2\}}} + \alpha_{T_{\{2\}}} \beta_{T_{\{1\}}})/2$, $d = d(T)$, $l = l(T) > 1$. So, by (6.2),

$$A_T(n) \sim \frac{2^{l(T)-1}}{2^{l(T)-1}-1} \left[ \frac{\alpha_{T_{\{1\}}} \alpha_{T_{\{2\}}}}{2} n^{d(T)} \right.$$

$$\left. + \left( \frac{\alpha_{T_{\{1\}}} \beta_{T_{\{2\}}} + \alpha_{T_{\{2\}}} \beta_{T_{\{1\}}}}{2} - \frac{d(T)\alpha_{T_{\{1\}}} \alpha_{T_{\{2\}}}/2}{2^{l(T)-1}} \right) n^{d(T)-1} \right] 2^{l(T)n}$$

So, using (6.4), we have

$$\beta_T = \frac{2^{l(T)-2}}{2^{l(T)-1}-1} \left( \alpha_{T_{\{1\}}} \beta_{T_{\{2\}}} + \alpha_{T_{\{2\}}} \beta_{T_{\{1\}}} - \frac{d(T)\alpha_{T_{\{1\}}} \alpha_{T_{\{2\}}}}{2^{l(T)-1}} \right)$$

$$= \frac{(\alpha_{T_{\{1\}}} \beta_{T_{\{2\}}} + \alpha_{T_{\{2\}}} \beta_{T_{\{1\}}})2^{l(T)-2} - d(T)\alpha_T}{2^{l(T)-1}-1},$$

as claimed.

Finally, if the root of $T$ has $r \geq 3$ lower covers $x_1, \ldots, x_r$ we can write (5.6) as

$$A_T(n) - 2A_T(n-1) = 2\sum_{j=1}^{r} \frac{1}{2} A_{T_{\{j\}}}(n)A_{T_{\{j\}^c}}(n-1)$$

$$+ \sum_{2 \leq |L| \leq r-2} A_{T_L}(n-1)A_{T_{L^c}}(n-1).$$

Terms in the first sum are of the form

$$(\alpha_{T_{\{j\}}} n^{d(T_{\{j\}})} + \beta_{T_{\{j\}}} n^{d(T_{\{j\}})-1})2^{l(T_{\{j\}})n}$$

$$\times (\alpha_{T_{\{j\}^c}}(n-1)^{d(T_{\{j\}^c})} + \beta_{T_{\{j\}^c}}(n-1)^{d(T_{\{j\}^c})-1})2^{l(T_{\{j\}^c})(n-1)}$$

$$\sim \frac{1}{2^{l(T_{\{j\}^c})}} \left[ \begin{array}{l} \alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}} n^{d(T)} \\ + \left( \alpha_{T_{\{j\}}} \beta_{T_{\{j\}^c}} + \alpha_{T_{\{j\}^c}} \beta_{T_{\{j\}}} - d(T_{\{j\}^c})\alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}} \right) n^{d(T)-1} \end{array} \right] 2^{l(T)n}$$

and terms in the second sum are of the form

$$(\alpha_{T_L}(n-1)^{d(T_L)} + \beta_{T_L}(n-1)^{d(T_L)-1})2^{l(T_L)(n-1)}$$

$$\times (\alpha_{T_{L^c}}(n-1)^{d(T_{L^c})} + \beta_{T_{L^c}}(n-1)^{d(T_{L^c})-1})2^{l(T_{L^c})(n-1)}$$

$$\sim \frac{1}{2^{l(T)}} \left[ \alpha_{T_L} \alpha_{T_{L^c}} n^{d(T)} + (\alpha_{T_L} \beta_{T_{L^c}} + \alpha_{T_{L^c}} \beta_{T_L} - d(T)\alpha_{T_L} \alpha_{T_{L^c}})n^{d(T)-1} \right] 2^{l(T)n},$$

since $l(T) = l(T_L) + l(T_{L^c})$ and $d(T) = d(T_L) + d(T_{L^c})$.

So, $A_T(n)$ satisfies a recurrence relation of the form in Lemma 6.1 with

$$\alpha = \sum_{j=1}^{r} \frac{\alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}}}{2^{l(T_{\{j\}^c})}} + \sum_{2 \leq |L| \leq r-2} \frac{\alpha_{T_L} \alpha_{T_{L^c}}}{2^{l(T)}},$$

$$\beta = \sum_{j=1}^{r} \frac{\alpha_{T_{\{j\}}} \beta_{T_{\{j\}^c}} + \alpha_{T_{\{j\}^c}} \beta_{T_{\{j\}}} - d(T_{\{j\}^c}) \alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}}}{2^{l(T_{\{j\}^c})}}$$

$$+ \sum_{2 \leq |L| \leq r-2} \frac{\alpha_{T_L} \beta_{T_{L^c}} + \alpha_{T_{L^c}} \beta_{T_L} - d(T) \alpha_{T_L} \alpha_{T_{L^c}}}{2^{l(T)}},$$

$d = d(T)$ and $l = l(T) > 1$. So, by (6.2), the coefficient $\beta_T$ of the term $n^{d(T)-1}$ in the leading polynomial $q_l$ of $A_T(n)$ is

$$\frac{2^{l(T)-1}}{2^{l(T)-1} - 1} \left( \begin{array}{l} \displaystyle\sum_{j=1}^{r} \frac{\alpha_{T_{\{j\}}} \beta_{T_{\{j\}^c}} + \alpha_{T_{\{j\}^c}} \beta_{T_{\{j\}}} - d(T_{\{j\}^c}) \alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}}}{2^{l(T_{\{j\}^c})}} \\[2em] + \displaystyle\sum_{2 \leq |L| \leq r-2} \frac{\alpha_{T_L} \beta_{T_{L^c}} + \alpha_{T_{L^c}} \beta_{T_L} - d(T) \alpha_{T_L} \alpha_{T_{L^c}}}{2^{l(T)}} \\[2em] - \displaystyle\frac{d(T)}{2^{l(T)-1} - 1} \left( \sum_{j=1}^{r} \frac{\alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}}}{2^{l(T_{\{j\}^c})}} + \sum_{2 \leq |L| \leq r-2} \frac{\alpha_{T_L} \alpha_{T_{L^c}}}{2^{l(T)}} \right) \end{array} \right).$$

Therefore, using (6.4), we have

$$\beta_T = \frac{\sum_{j=1}^{r} (\alpha_{T_{\{j\}}} \beta_{T_{\{j\}^c}} + \alpha_{T_{\{j\}^c}} \beta_{T_{\{j\}}} - d(T_{\{j\}^c}) \alpha_{T_{\{j\}}} \alpha_{T_{\{j\}^c}}) 2^{l(T_{\{j\}})-1}}{2^{l(T)-1} - 1}$$

$$+ \frac{\sum_{2 \leq |L| \leq r-2} (\alpha_{T_L} \beta_{T_{L^c}} + \alpha_{T_{L^c}} \beta_{T_L} - d(T) \alpha_{T_L} \alpha_{T_{L^c}}) 2^{-1} - d(T) \alpha_T}{2^{l(T)-1} - 1},$$

as claimed.                                                                                    $\square$

## 6.2  Typical embeddings of $T$ into $T^n$

We have that, for $T$ a tree with $|T| > 1$, $A_T(n) \sim \alpha_T n^{d(T)} 2^{l(T)n}$, for $\alpha_T$ some constant that can be found. Let us give an informal description of a "typical" embedding of $T$ into $T^n$, giving an alternative method of seeing at least the lower bound $A_T(n) = \Omega(n^{d(T)} 2^{l(T)n})$. For any tree $T$, call the elements counted by $d(T)$ *lower bead elements* of $T$. So, a lower bead element of $T$ is an element $x$ such that

$D[x]$ is a chain, and $x$ is not a leaf or the root. Call an element which has more than one lower cover a *branching element* of $T$. Call the remaining elements of $T$ different from its leaves the *upper bead elements* of $T$. These are elements $x$ which have only one lower cover, but $D[x]$ is not a chain. Therefore, upper bead elements only occur on a chain above a branching element. Note that, depending on the tree $T$, the root can be either a branching element or an upper bead element.

So, if $T$ is a chain, then $T$ has a root and one leaf, joined by a chain of $d(T)$ lower bead elements. Otherwise, for $l(T) > 1$, the tree $T$ has a root, the root and the branching elements are joined by (possibly empty) chains of upper bead elements, and some branching elements are joined by (possibly empty) chains of lower bead elements (of which there are $d(T)$) to the $l(T)$ leaves.

To see that $A_T(n) = \Omega(n^{d(T)} 2^{l(T)n})$, first consider $T$ a chain. We count the embeddings that map the root of $T$ to $1_n$ and the leaf of $T$ to some leaf of $T^n$. We have $2^{n-1}$ choices for where to map the leaf. Once we have fixed the leaf of $T^n$, this defines a path from $1_n$ to the leaf of $T^n$. This gives a choice of $n-2$ elements of $T^n$ into which we can map the $d(T)$ lower bead elements of $T$. So, asymptotically we have $\Theta(n^{d(T)})$ choices for where to map the $d(T)$ lower bead elements. Therefore $A_T(n) = \Omega(n^{d(T)} 2^n)$, and since $l(T) = 1$ we have that $A_T(n) = \Omega(n^{d(T)} 2^{l(T)n})$ for $T$ a chain.

For $T$ not a chain, so there exist branching elements of $T$, let $\phi$ be some embedding which maps the root of $T$ to $1_n$, and maps the branching elements of $T$ to as high a level of $T^n$ as possible. Consider, for large $n$, the number of embeddings of $T$ into $T^n$ that agree with this fixed $\phi$ on the root, branching elements and upper bead elements. Let us only consider those embeddings which map the leaves of $T$ to the leaves of $T^n$. If a leaf $y$ is joined to a branching element $x$ by a chain of lower bead elements, then note that $\phi(x)$ is a fixed distance from the root of $T^n$, so that $\phi(x)$ is in level $n - k_x$ of $T^n$, where $k_x$ is a constant independent of $n$. So, given $\phi$, the leaf $y$ can be mapped to $2^{-k_x} 2^{n-1}$ leaves in $T^n$. So, the total number of choices for all the

leaves is asymptotically $\Theta(2^{l(T)n})$. (The over-counting due to the possibility that two leaves that are below the same branching point are mapped to the same leaf of $T^n$ is negligible for large $n$.) It remains to choose where to map the lower bead elements. However, in a similar way to the case where $T$ is a chain, a lower bead element on the chain between the branching point $x$ and the leaf $y$ must be mapped to an element on the path between the images of $x$ and $y$. Since $x$ is mapped to level $n - k_x$, and $y$ to a leaf, the path has $n - k_x - 2$ elements, with $k_x$ independent of $n$. Since there are $d(T)$ lower bead elements, we have asymptotically $\Theta(n^{d(T)})$ choices for where to map the lower bead elements. (Again, this is an over-count due to the possibility that two lower bead elements that are below the same branching element but above different leaves are mapped to the same element of $T^n$. However, this is negligible because typically the lower bead elements will not be mapped within $O(1)$ of a branching element.) So, the number of embeddings that agree with $\phi$ is asymptotically $\Omega(n^{d(T)}2^{l(T)n})$, and we have $A_T(n) = \Omega(n^{d(T)}2^{l(T)n})$ for $T$ not a chain.

By Lemma 6.1 we also have the asymptotic behaviour of $C_T(n)$, given in the following corollary.

**Corollary 6.3.** *For any tree $T$ with $l(T) = 1$ the number of embeddings of $T$ into $T^n$ is asymptotically*

$$C_T(n) \sim \frac{\alpha_T}{d(T) + 1} n^{d(T)+1} 2^n$$

*and if $d(T) > 0$ then*

$$C_T(n) \sim \left( \frac{\alpha_T}{d(T) + 1} n^{d(T)+1} + \left( \frac{\beta_T}{d(T)} + \frac{\alpha_T}{2} \right) n^{d(T)} \right) 2^n.$$

*For any tree with $l(T) > 1$ the number of embeddings of $T$ into $T^n$ is asymptotically*

$$C_T(n) \sim \frac{2^{l(T)-1}}{2^{l(T)-1} - 1} \alpha_T n^{d(T)} 2^n$$

*and if $d(T) > 0$ then*

$$C_T(n) \sim \frac{2^{l(T)-1}}{2^{l(T)-1} - 1} \left( \alpha_T n^{d(T)} + \left( \beta_T - \frac{d(T)\alpha_T}{2^{l(T)-1} - 1} \right) n^{d(T)-1} \right) 2^{l(T)n}.$$

**Proof.** We have that $A_T(n) \sim \alpha_T n^{d(T)} 2^{l(T)n}$, and if $d(T) > 0$ then $A_T(n) \sim (\alpha_T n^{d(T)} + \beta_T n^{d(T)-1}) 2^{l(T)n}$. So $C_T(n)$ satisfies the recurrence relation (5.1) which is of the form in Lemma 6.1. Applying Lemma 6.1 with $\alpha = \alpha_T$ and $\beta = \beta_T$ gives the result. $\qquad\square$

This tells us that for a tree $T$ not a chain, a typical embedding of $T$ into $T^n$ maps the leaves of $T$ to the low levels of $T^n$, the branching points and upper bead elements of $T$ to the high levels of $T^n$, and the lower bead elements of $T$ will be mapped to elements spread roughly evenly along the paths in $T^n$ defined by the images of branching elements and leaves of $T$, as explained earlier. There are $\Theta(n^{d(T)} 2^{l(T)n})$ of these embeddings.

For $T$ a chain, a typical embedding maps the leaf of $T$ to a low level of $T^n$, and the remaining elements of $T$ are mapped to elements spread roughly evenly on the path from $1_n$ to image of the leaf in $T^n$. Here the root is not necessarily mapped to $1_n$, and the root can be thought of as a lower bead element, so there are $d(T) + 1$ elements to position on this path. So, we get $\Theta(n^{d(T)+1} 2^n)$ of these embeddings.

## 6.3 Asymptotics of the ratio $A_T(n)/C_T(n)$

In [18], Kubicki, Lehel and Morayne proved that $\lim_{n\to\infty} \frac{A_{T_1}(n)}{B_{T_1}(n)} \le \lim_{n\to\infty} \frac{A_{T_2}(n)}{B_{T_2}(n)}$, where $B_T(n)$ is the number of embeddings $\phi$ of $T$ into $T^n$ with $\phi(1_T) \ne \phi(1_n)$, by showing that $\lim_{n\to\infty} A_T(n)/B_T(n) = 2^{l(T)-1} - 1$ (Proposition 2.3 in [18]). Here, using Theorem 6.2 and Corollary 6.3 we have

$$\lim_{n\to\infty} \frac{A_T(n)}{C_T(n)} = \frac{2^{l(T)-1} - 1}{2^{l(T)-1}}$$

which is equivalent to Proposition 2.3 in [18], since $B_T(n) = C_T(n) - A_T(n)$. This tells us that for trees $T_1, T_2$ with $l(T_1) < l(T_2)$ there exists some $n_0$ such that $A_{T_1}(n)/C_{T_1}(n) < A_{T_2}(n)/C_{T_2}(n)$ for all $n \ge n_0$. Here, we show that there exist trees $T_1 \subseteq T_2$, with $l(T_1) = l(T_2)$, with the inequality the other way round. That is, there

is an $n_0$ such that $A_{T_1}(n)/C_{T_1}(n) > A_{T_2}(n)/C_{T_2}(n)$ for all $n \geq n_0$. All such pairs $T_1, T_2$ are counterexamples to the conjecture, for all $n \geq n_0$.

**Theorem 6.4.** *For any tree $T$ with $l(T) > 1$ and $d(T) > 0$, we have*

$$\frac{A_T(n)}{C_T(n)} = 1 - \frac{1}{2^{l(T)-1}}\left(1 - \frac{d(T)}{n} + \frac{\binom{d(T)}{2}}{n^2} + \frac{b_T}{n^2}\right) + o(n^{-2}) \qquad (6.6)$$

*where*

$$b_T = \frac{\beta_T}{\alpha_T} - \frac{d(T)}{2^{l(T)-1}-1}. \qquad (6.7)$$

*For any tree $T$ with $l(T) > 1$ and $d(T) = 0$, we have*

$$\frac{A_T(n)}{C_T(n)} = 1 - \frac{1}{2^{l(T)-1}} + o(n^{-1}). \qquad (6.8)$$

*For any tree $T$ with $l(T) = 1$, we have*

$$\frac{A_T(n)}{C_T(n)} = \frac{d(T)+1}{n} + o(n^{-1}). \qquad (6.9)$$

**Proof.** Let $T$ be a tree with $l(T) > 1$ and $d(T) > 0$. By (5.1) it is sufficient to work with the ratio $C_T(n-1)/C_T(n)$. By Theorem 5.1 we have that $C_T(n) = \sum_{j=0}^{l(T)} q_j(n)2^{jn}$ and by Corollary 6.3 we have that

$$q_l(n) \sim \frac{2^{l(T)-1}}{2^{l(T)-1}-1}\left(\alpha_T n^d(T) + \left(\beta_T - \frac{\alpha_T d(T)}{2^{l(T)-1}-1}\right)n^{d(T)-1}\right).$$

So,

$$C_T(n) = 2^{l(T)n}a_T(n^{d(T)} + b_T n^{d(T)-1} + c_T n^{d(T)-2} + o(n^{d(T)-2}))$$

where

$$a_T = \frac{2^{l(T)-1}}{2^{l(T)-1}-1}\alpha_T, \qquad\qquad b_T = \frac{\beta_T}{\alpha_T} - \frac{d(T)}{2^{l(T)-1}-1}$$

and $c_T$ is an unspecified constant. Note that this equation is true for $d \geq 2$, and

can be made true for $d = 1$ by setting $c_T$ to 0. We have

$$
\frac{C_T(n-1)}{C_T(n)} = \frac{2^{l(T)(n-1)}a_T((n-1)^{d(T)} + b_T(n-1)^{d(T)-1} + c_T(n-1)^{d(T)-2} + o(n^{d(T)-2}))}{2^{l(T)n}a_T(n^{d(T)} + b_T n^{d(T)-1} + c_T n^{d(T)-2} + o(n^{d(T)-2}))}
$$

$$
= \frac{1}{2^{l(T)}} \frac{(1 - 1/n)^{d(T)} + \frac{b_T}{n}(1 - 1/n)^{d(T)-1} + \frac{c_T}{n^2}(1 - 1/n)^{d(T)-2} + o(n^{-2})}{1 + b_T/n + c_T/n^2 + o(n^{-2})}
$$

$$
= \frac{1}{2^{l(T)}} \left( 1 - \frac{d(T)}{n} + \frac{\binom{d(T)}{2}}{n^2} + \frac{b_T}{n} - \frac{b_T(d(T) - 1)}{n^2} + \frac{c_T}{n^2} + o(n^{-2}) \right)
$$

$$
\times \left( 1 - \frac{b_T}{n} - \frac{c_T}{n^2} + \frac{b_T^2}{n^2} + o(n^{-2}) \right)
$$

$$
= \frac{1}{2^{l(T)}} \left( 1 - \frac{d(T)}{n} + \frac{\binom{d(T)}{2}}{n^2} + \frac{b_T}{n^2} + o(n^{-2}) \right)
$$

and, using (5.1), we have

$$
\frac{A_T(n)}{C_T(n)} = 1 - \frac{1}{2^{l(T)-1}} \left( 1 - \frac{d(T)}{n} + \frac{\binom{d(T)}{2}}{n^2} + \frac{b_T}{n^2} + o(n^{-2}) \right)
$$

as required.

Now, suppose $l(T) > 1$ and $d(T) = 0$. So, $C_T(n) = a_T 2^{l(T)n} + \sum_{j=0}^{l(T)-1} q_j(n)2^{jn}$. That is, $C_T(n) = a_T 2^{l(T)n}(1 + O(g(n)2^{-n}))$ for some polynomial $g(n)$. So, it is certainly true that $C_T(n) = a_T 2^{l(T)n}(1 + o(n^{-1}))$ and

$$
\frac{C_T(n-1)}{C_T(n)} = \frac{1}{2^{l(T)}}(1 + o(n^{-1}))
$$

which by (5.1) gives the required result.

If $l(T) = 1$, then $A_T(n) = 2^n \alpha_T(n^{d(T)} + o(n^{d(T)}))$ and $C_T(n) = 2^n \frac{\alpha_T}{d(T)+1}(n^{d(T)+1} + o(n^{d(T)+1}))$. So,

$$
\frac{A_T(n)}{C_T(n)} = \frac{2^n \alpha_T(n^{d(T)} + o(n^{d(T)}))}{2^n \frac{\alpha_T}{d(T)+1}(n^{d(T)+1} + o(n^{d(T)+1}))} = \frac{d(T) + 1}{n}(1 + o(1)). \qquad \square
$$

**Corollary 6.5.** *For any two trees $T_1, T_2$, if either*

*(i) $l(T_1) > l(T_2)$, or*

*(ii) $l(T_1) = l(T_2)$ and $d(T_1) > d(T_2)$, or*

*(iii) $l(T_1) = l(T_2)$, $d(T_1) = d(T_2) > 0$ and $\alpha_{T_1}/\beta_{T_1} > \alpha_{T_2}/\beta_{T_2}$,*

*then there exists an integer $n_0$ such that*

$$\frac{A_{T_1}(n)}{C_{T_1}(n)} > \frac{A_{T_2}(n)}{C_{T_2}(n)}$$

*for all $n \geq n_0$.*

**Proof.** (i) If $l(T_1) > l(T_2)$ then we can just compare the limits of the ratios $A_{T_1}(n)/C_{T_1}(n)$ and $A_{T_2}(n)/C_{T_2}(n)$. By Theorem 6.4 (or from [18]) we have that

$$\lim_{n \to \infty} \frac{A_T(n)}{C_T(n)} = 1 - \frac{1}{2^{l(T)-1}}.$$

Note that this also holds for trees $T$ with $l(T) = 1$. Since the limit is increasing in $l(T)$ the result follows.

(ii) If $l(T_1) = l(T_2)$ and $d(T_1) > d(T_2)$ there are two cases to consider. If $l(T_1) = l(T_2) = 1$ then using equation (6.9) from Theorem 6.4 we have that

$$\frac{A_{T_1}(n)}{C_{T_1}(n)} = \frac{d(T_1)+1}{n} + o(n^{-1}) \qquad \frac{A_{T_2}(n)}{C_{T_2}(n)} = \frac{d(T_2)+1}{n} + o(n^{-1})$$

and since $d(T_1) > d(T_2)$ there exists an $n_0$ such that $A_{T_1}(n)/C_{T_1}(n) > A_{T_2}(n)/C_{T_2}(n)$ for all $n \geq n_0$.

If $l(T_1) = l(T_2) > 1$ then using equation (6.6) from Theorem 6.4, and considering only terms up to $n^{-1}$ we have

$$\frac{A_{T_1}(n)}{C_{T_1}(n)} = 1 - \frac{1}{2^{l(T_1)-1}}\left(1 - \frac{d(T_1)}{n}\right) + o(n^{-1}),$$

$$\frac{A_{T_2}(n)}{C_{T_2}(n)} = 1 - \frac{1}{2^{l(T_2)-1}}\left(1 - \frac{d(T_2)}{n}\right) + o(n^{-1}).$$

This is also true for $d(T_2) = 0$ by equation (6.8). Since $l(T_1) = l(T_2)$ and $d(T_1) > d(T_2)$ there exists an $n_0$ such that $A_{T_1}(n)/C_{T_1}(n) > A_{T_2}(n)/C_{T_2}(n)$ for all $n \geq n_0$.

(iii) If $l(T_1) = l(T_2)$ and $d(T_1) = d(T_2) > 0$ and $\alpha_{T_1}/\beta_{T_1} > \alpha_{T_2}/\beta_{T_2}$, we first note that $l(T_1)$ cannot be equal to 1. (If $l(T_1) = l(T_2) = 1$ then $d(T_1) = d(T_2)$ implies that $T_1$ and $T_2$ are the same tree, the $(d+2)$-element chain.) So we have $l(T_1) = l(T_2) > 1$ and using equation (6.6), we see that $A_{T_1}(n)/C_{T_1}(n)$ and $A_{T_2}(n)/C_{T_2}(n)$ differ only

in the $n^{-2}$ term and in terms of lower order. Therefore, it is enough to show that $b_{T_1} < b_{T_2}$. But this follows immediately from the inequality $\alpha_{T_1}/\beta_{T_1} > \alpha_{T_2}/\beta_{T_2}$ and (6.7). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

# 6.4 A family of counterexamples to Conjecture 4.4 for arbitrarily large $n$

Corollary 6.5 provides a simple method for comparing the asymptotics of the ratios $A_{T_1}(n)/C_{T_1}(n)$ and $A_{T_2}(n)/C_{T_2}(n)$. Firstly, we compare the number of leaves of the two trees, the tree with more leaves being the tree with the asymptotically larger ratio $A/C$. If the trees have the same number of leaves, then we compare the values of $d(T_1)$ and $d(T_2)$; the tree with the larger $d$ has the asymptotically larger ratio $A/C$. Both the number of leaves, $l(T)$, and $d(T)$ are very easily obtained from the Hasse diagram of the tree. If both of these are the same for the two trees, then we need to compare the ratios $\alpha_{T_1}/\beta_{T_1}$ and $\alpha_{T_2}/\beta_{T_2}$. The tree with the larger ratio $\alpha/\beta$ has the asymptotically larger ratio $A/C$. This comparison involves rather more calculation, using the algorithm provided by Theorem 6.2. These calculations can be simplified if the two trees have a very similar structure, for example, as we will see later, if the trees are identical except for the addition of one element to one of the trees.

Corollary 6.5 also guides our search for more counterexamples to the conjecture of Kubicki, Lehel and Morayne. The counterexample given in Section 5.2 has two important properties, namely that $l(T_1) = l(T_2)$ and $d(T_1) = d(T_2)$. That this is a necessary condition for a pair of trees to be an asymptotic counterexample follows from Corollary 6.5. Since we are only considering trees $T_1 \subseteq T_2$ we must have $l(T_1) \leq l(T_2)$. But we are looking for trees $T_1, T_2$ where the ratio $A/C$ is asymptotically larger for $T_1$ than for $T_2$, so we need to look at trees with $l(T_1) = l(T_2)$. If $T_1 \subseteq T_2$ and the trees have the same number of leaves we must have $d(T_1) \leq d(T_2)$. (Each

element counted by $d(T_1)$ must also be counted by $d(T_2)$ otherwise $T_2$ would have more leaves than $T_1$.) So, to find our counterexamples we need to look at trees with $d(T_1) = d(T_2)$.

The following theorem gives an infinite family of pairs of trees which form counterexamples. We do not claim, or believe, that this is the only way to construct counterexamples. However, the construction is relatively simple, which makes the calculations much more manageable. Also, there are many ternary tree pairs in this family, including the counterexample given in Section 5.2, which shows that the conjecture does not just fail for trees with high branching numbers.

**Theorem 6.6.** *Let $T$ be a tree whose root $x$ has three lower covers $x_1, x_2, x_3$, and let $T'$ be formed from $T$ by adding a new element $y$ below $x$ and above $x_2$ and $x_3$ (see Figure 6.1). If $d(T) > 0$ and $d(D[y]) = 0$, then there exists $n_0$ such that $A_T(n)/C_T(n) > A_{T'}(n)/C_{T'}(n)$ for all $n \geq n_0$.*
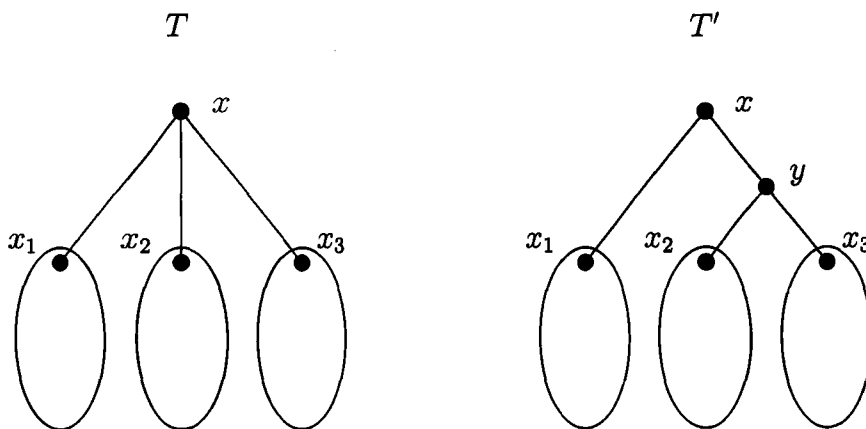


Figure 6.1: General counterexample for $d(T) > 0, d(D[y]) = 0$

**Proof.** We have $l(T) = l(T')$ and $d(T) = d(T') > 0$ so by Corollary 6.5 it is enough to show that $\alpha_T \beta_{T'} > \alpha_{T'} \beta_T$. We use equations (6.4) and (6.5) to express these $\alpha$ and $\beta$ in terms of some other $\alpha_S$ and $\beta_S$ for common subtrees $S$ of $T$ and $T'$. As before, for $L \subseteq [3]$ write $T_L$ for the subtree formed from $T$ by removing the elements in $D[x_j]$ for each $j \in L^c$. Write $T'_{\{1\}}$ for the subtree formed from $T'$ by removing elements in

$D[y]$ and write $T_y$ for the subtree formed by removing elements in $D[x_1]$. We have

that $T_{\{1\}} = T'_{\{1\}}$ and $T_{\{2,3\}} \cong D[y]$. By the assumption that $d(D[y]) = 0$ we have

that $d(T) = d(T') = d(T_{\{1,2\}}) = d(T_{\{1,3\}}) = d(T_{\{1\}})$, and we denote this common

value by $d$. We also have that $d(T_{\{2\}}) = d(T_{\{3\}}) = d(T_{\{y\}}) = d(D[y]) = 0$. For ease

of notation, we write $l$ for the common value $l(T) = l(T')$, write $l_1$ for $l(T_{\{1\}})$, $l_{12}$

for $l(T_{\{1,2\}})$, etc., and we use a similar notation for $\alpha$ and $\beta$. For example, writing

$\alpha_1$ for $\alpha_{T_{\{1\}}}$.

Using equation (6.5) to find $\beta_T$ and $\beta_{T'}$, we have

$$\beta_T = \frac{\alpha_{23}\beta_1 2^{l_1-1} + (\alpha_2\beta_{13} - d\alpha_2\alpha_{13})2^{l_2-1} + (\alpha_3\beta_{12} - d\alpha_3\alpha_{12})2^{l_3-1} - d\alpha_T}{2^{l-1} - 1},$$

$$\beta_{T'} = \frac{\alpha_y\beta_1 2^{l-2} - d\alpha_{T'}}{2^{l-1} - 1}$$

so

$$\alpha_T\beta_{T'} - \alpha_{T'}\beta_T = \frac{1}{2^{l-1} - 1}\left[\begin{array}{l} \alpha_T\alpha_y\beta_1 2^{l-2} - \alpha_{T'}\left(\alpha_{23}\beta_1 2^{l_1-1} + (\alpha_2\beta_{13} - d\alpha_2\alpha_{13})2^{l_2-1}\right) \\ - \alpha_{T'}(\alpha_3\beta_{12} - d\alpha_3\alpha_{12})2^{l_3-1} \end{array}\right]$$

and using (6.4) to find $\alpha_T$ and $\alpha_{T'}$ we have

$$\alpha_T = \frac{\alpha_1\alpha_{23}2^{l_1-1} + \alpha_2\alpha_{13}2^{l_2-1} + \alpha_3\alpha_{12}2^{l_3-1}}{2^{l-1} - 1},$$

$$\alpha_{T'} = \frac{\alpha_1\alpha_y 2^{l-2}}{2^{l-1} - 1}.$$

This gives

$$\frac{(\alpha_T\beta_{T'} - \alpha_{T'}\beta_T)(2^{l-1} - 1)^2}{\alpha_y 2^{l-2}} = \left(\alpha_2\alpha_{13}2^{l_2-1} + \alpha_3\alpha_{12}2^{l_3-1}\right)\beta_1$$

$$- \alpha_1\left(\begin{array}{l} (\alpha_2\beta_{13} - d\alpha_2\alpha_{13})2^{l_2-1} \\ + (\alpha_3\beta_{12} - d\alpha_3\alpha_{12})2^{l_3-1} \end{array}\right)$$

$$= 2^{l_2-1}\alpha_2(\alpha_{13}\beta_1 - \alpha_1\beta_{13} + d\alpha_1\alpha_{13})$$

$$+ 2^{l_3-1}\alpha_3(\alpha_{12}\beta_1 - \alpha_1\beta_{12} + d\alpha_1\alpha_{12}).$$

Finally, we have

$$\beta_{13} = \frac{\beta_1\alpha_3 2^{l_{13}-2} - d\alpha_{13}}{2^{l_{13}-1} - 1} \quad \text{and} \quad \alpha_{13} = \frac{\alpha_1\alpha_3 2^{l_{13}-2}}{2^{l_{13}-1} - 1}$$

so

$$\alpha_{13}\beta_1 - \alpha_1\beta_{13} + d\alpha_1\alpha_{13} = \frac{\alpha_1\alpha_3 2^{l_{13}-2}}{2^{l_{13}-1}-1}\beta_1 - \alpha_1\frac{\beta_1\alpha_3 2^{l_{13}-2} - d\alpha_{13}}{2^{l_{13}-1}-1} + d\alpha_1\alpha_{13}$$
$$= \frac{d\alpha_1\alpha_{13}2^{l_{13}-1}}{2^{l_{13}-1}-1}$$

and similarly

$$\alpha_{12}\beta_1 - \alpha_1\beta_{12} + d\alpha_1\alpha_{12} = \frac{d\alpha_1\alpha_{12}2^{l_{12}-1}}{2^{l_{12}-1}-1}.$$

Therefore

$$\alpha_T\beta_{T'} - \alpha_{T'}\beta_T = \frac{\alpha_y 2^{l-2}}{(2^{l-1}-1)^2}\left[2^{l_2-1}\alpha_2\frac{d\alpha_1\alpha_{13}2^{l_{13}-1}}{2^{l_{13}-1}-1} + 2^{l_3-1}\alpha_3\frac{d\alpha_1\alpha_{12}2^{l_{12}-1}}{2^{l_{12}-1}-1}\right]$$
$$= \frac{\alpha_y\left(2^{l-2}\right)^2 d\alpha_1}{(2^{l-1}-1)^2}\left[\frac{\alpha_2\alpha_{13}}{2^{l_{13}-1}-1} + \frac{\alpha_3\alpha_{12}}{2^{l_{12}-1}-1}\right]$$
$$= \frac{d\alpha_{T'}2^{l-2}}{2^{l-1}-1}\left[\frac{\alpha_2\alpha_{13}}{2^{l_{13}-1}-1} + \frac{\alpha_3\alpha_{12}}{2^{l_{12}-1}-1}\right]$$

and since $\alpha_S > 0$ for all trees $S$, we have $\alpha_T\beta_{T'} - \alpha_{T'}\beta_T > 0$ as required.    $\square$

# Chapter 7

# Results for the complete $p$-ary tree

We can generalise the results of Chapters 5 and 6 to embeddings of trees into the complete $p$-ary tree. The aim of this chapter is to show that a most accommodating version of Conjecture 4.4 is still false, namely that even for embeddings into $p$-ary trees with $p > 2$, there exists a pair of ternary trees $T_1, T_2$ with $T_1$ a subposet of $T_2$ such that

$$\frac{A_{T_1}^{(p)}(n)}{C_{T_1}^{(p)}(n)} > \frac{A_{T_2}^{(p)}(n)}{C_{T_2}^{(p)}(n)}$$

for all $n \geq n_0$, for some $n_0$. This means that, even in this setting, the restriction on $T_1, T_2$ being binary cannot be removed.

We present some of the results in this chapter without proof, since they are the exact analogues of the results for the particular case $p = 2$.

## 7.1 Recurrence relations for $A_T^{(p)}(n)$ and $C_T^{(p)}(n)$

Let $t_1, \ldots, t_p$ be the $p$ lower covers of the root of $T_p^n$, and let $T_{p,i}^n$ be the elements of $T_p^n$ below or equal to $t_i$, for $i = 1 \ldots, p$. The subtrees $T_{p,i}^n$ are copies of $T_p^{n-1}$.

So, the recurrence relation corresponding to (5.1) is

$$C_T^{(p)}(n) - pC_T^{(p)}(n-1) = A_T^{(p)}(n). \tag{7.1}$$

Let $T$ be a tree and suppose its root $1_T$ has $r$ lower covers $x_1, \ldots, x_r$. We are now interested in the partition of $\{x_1, \ldots, x_r\}$ defined by which of the $p$ subtrees $T_{p,1}^n, \ldots, T_{p,p}^n$ an element $x_i$ is mapped to. For $L \subseteq [r]$ write $A_{T_L}^{(p)-}(n)$ for the number of embeddings of $T_L$ into $T_p^n$ that map the root $1_T$ of $T_L$ to $1_n$ and map $x_j$ into $T_{p,1}^n$ for each $j \in L$. As for the complete binary tree, this number is the same as the number of embeddings of $T_L$ into $T_p^n$ that map $1_T$ to $1_n$ and map $x_j$ into $T_{p,i}^n$ for each $j \in L$, for *any* $i = 1, \ldots, p$.

Write $(L_1, \ldots, L_p) \vdash [r]$ to mean that the sets $L_1, \ldots, L_p$ form a partition of $[r]$, so that $\bigcup_{i=1}^p L_i = [r]$, $L_i \cap L_j = \emptyset$ for all $i \neq j$. We have

$$A_T^{(p)}(n) = \sum_{\substack{L_1, \ldots, L_p: \\ (L_1, \ldots, L_p) \vdash [r]}} \prod_{i=1}^p A_{T_{L_i}}^{(p)-}(n) \tag{7.2}$$

corresponding to equation (5.2),

$$A_{T_L}^{(p)-}(n) = \begin{cases} 1 & \text{if } L = \emptyset \\ C_{D[x_j]}^{(p)}(n-1) & \text{if } L = \{j\} \\ A_{T_L}^{(p)}(n-1) & \text{otherwise} \end{cases} \tag{7.3}$$

for all $i = 1, \ldots, p$, corresponding to equation (5.3), and

$$A_{T_L}^{(p)-}(n) = \frac{A_{T_L}^{(p)}(n)}{p} \tag{7.4}$$

for $L = \{j\}$, corresponding to equation (5.4).

We have the following result for embeddings into $T_p^n$, analogous to the complete binary tree case.

**Theorem 7.1.** *Let $p$ be an integer with $p \geq 2$. For any tree $T$, the number of embeddings of $T$ into $T_p^n$ is of the form*

$$C_T^{(p)}(n) = \sum_{j=0}^{l(T)} g_j(n) p^{jn},$$

*where each $g_j$ is a polynomial.*

*For T the 1-element tree, the number of these embeddings that map the root of T to $1_n$, $A_T^{(p)}(n)$, is equal to 1. Otherwise, for T with $|T| > 1$, the number is of the form*

$$A_T^{(p)}(n) = \sum_{j=0}^{l(T)} q_j(n) p^{jn},$$

*where each $q_j$ is a polynomial.*

To prove this, we again use some results on recurrence relations; here we need the following generalisation of Lemma 5.2.

**Lemma 7.2.** *Let p be an integer with $p \geq 2$, and suppose l is some fixed positive integer. Then the solution to the equation*

$$y_n - p y_{n-1} = \sum_{j=0}^{l} f_j(n) p^{jn}, \quad y_1 = 0, \tag{7.5}$$

*where each $f_j$ is a polynomial, is*

$$y_n = \sum_{j=0}^{l} g_j(n) p^{jn}$$

*where each $g_j$ is a polynomial. Furthermore, for $j \neq 1$, the polynomial $g_j$ is the unique polynomial satisfying the identity*

$$g_j(n) - p^{1-j} g_j(n-1) = f_j(n),$$

*and $g_1$ satisfies the identity*

$$g_1(n) - g_1(n-1) = f_1(n),$$

*where the constant term of $g_1$ is given by*

$$\sum_{j=0}^{l} g_j(1) p^j = 0$$

**Proof.** The proof method exactly follows that of Lemma 5.2, with 2 replaced by $p$.      □

**Proof of Theorem 7.1.** The proof follows the same method as for Theorem 5.1, but with slight modifications because of the difference between equations (5.2) and (7.2).

We include the case of $T$ being a 1-element set for completeness. In this case, we see immediately that there are $(p^n - 1)/(p-1)$ embeddings of $T$ into $T_p^n$, which is exactly the number of elements in $T_p^n$. Also, only one of these embeddings maps the root of $T$ to $1_n$. So, $A_T^{(p)}(n) = 1$ as claimed, and $C_T^{(p)}(n) = (p^n - 1)/(p-1)$ is of the required form.

For $|T| \geq 2$, we simultaneously prove that $A_T^{(p)}(n)$ and $C_T^{(p)}(n)$ are of the required form by induction on the size of $T$. We shall make use of Lemma 7.2 to solve recurrence relations for $A_T^{(p)}(n)$ and $C_T^{(p)}(n)$. We use induction to show that the recurrence is of the form of equation (7.5), and since we will only be considering trees with $|T| \geq 2$ we have the initial conditions $A_T^{(p)}(1) = 0, C_T^{(p)}(1) = 0$ as in (7.5).

For $|T| = 2$ the only tree is the 2-element chain, which has one leaf. Since the root $1_T$ has only one lower cover $x_1$, say, we have $r = 1$ in (7.2). The only partitions of the set $[1] = \{1\}$ are those with $L_i = \{1\}$ for exactly one $i$, and $L_j = \emptyset$ for all $j \neq i$, a total of $p$ different partitions. Using (7.3) with $L_i = \{1\}$ or $L_i = \emptyset$, equation (7.2) becomes

$$A_T^{(p)}(n) = pC_{\{x_1\}}^{(p)}(n - 1).$$

(Compare this with the binary case, where $p = 2$.) We have shown earlier that $C_{\{x_1\}}^{(p)}(n) = (p^n - 1)/(p-1)$. Therefore $A_T^{(p)}(n) = (p^n - p)/(p-1)$ which is of the required form, since $l(T) = 1$, $q_0(n) = -p/(p-1)$ and $q_1(n) = 1/(p-1)$. Using (7.1) and Lemma 7.2 we have that

$$C_T^{(p)}(n) = \frac{p^n((p-1)n - p) + p}{(p-1)^2}$$

which is of the required form, since $l(T) = 1$, $g_0(n) = p/(p-1)^2$ and $g_1(n) = ((p-1)n - p)/(p-1)^2$.

Suppose the result is true for all $T$ with $|T| < k$ and let $T$ be any tree with

$|T| = k$. As before, there are two cases, depending on whether the root of $T$ has exactly one lower cover. If the root has exactly one lower cover, $x_1$, equation (7.2) reduces, in a similar way to the base case, to

$$A_T^{(p)}(n) = pC_{D[x_1]}^{(p)}(n-1).$$

Applying the inductive hypothesis to $D[x_1]$, a tree with $l(D[x_1]) = l(T)$ leaves, gives

$$C_{D[x_1]}^{(p)}(n) = \sum_{j=0}^{l(T)} g_j(n)p^{jn},$$

where $g_j$ are polynomials. Therefore,

$$A_T^{(p)}(n) = p\sum_{j=0}^{l(T)} g_j(n-1)p^{j(n-1)} = \sum_{j=0}^{l(T)} q_j(n)p^{jn}$$

where $q_j$ are polynomials.

If the root of $T$ has $r > 1$ lower covers $x_1, \ldots, x_r$, then $[r]$ has exactly $p$ partitions with $L_i = [r]$ for some $i$, and $L_j = \emptyset$ for all $j \neq i$. All other partitions have $L_i \neq [r]$ for all $i = 1, \ldots, p$. So equation (7.2) becomes

$$A_T^{(p)}(n) = pA_T^{(p)}(n-1) + \sum_{\substack{L_1,\ldots,L_p: \\ (L_1,\ldots,L_p)\vdash[r], \\ L_i \neq [r]}} \prod_{i=1}^{p} A_{T_{L_i}}^{(p)-}(n),$$

or, equivalently,

$$A_T^{(p)}(n) - pA_T^{(p)}(n-1) = \sum_{\substack{L_1,\ldots,L_p: \\ (L_1,\ldots,L_p)\vdash[r], \\ L_i \neq [r]}} \prod_{i=1}^{p} A_{T_{L_i}}^{(p)-}(n). \tag{7.6}$$

It remains to show that this equation is a recurrence relation of the form of equation (7.5), as follows.

Each term $A_{T_{L_i}}^{(p)-}(n)$ is either 1, $A_{T_{L_i}}^{(p)}(n)/p$ for some $L_i$ a singleton, or $A_{T_{L_i}}^{(p)}(n-1)$ for some $L_i$ not a singleton. Since all trees $T_{L_i}$ have fewer elements than $T$ (by the condition that $L_i \neq [r]$) we can apply our inductive hypothesis and we have $A_{T_{L_i}}^{(p)}(n) = \sum_{j=0}^{l(T_{L_i})} q_j(n)p^{jn}$ for polynomials $q_j$. This means each term $A_{T_{L_i}}^{(p)-}(n)$ is

either 1, or of the form $\sum_{j=0}^{l(T_{L_i})} r_j(n)p^{jn}$ for polynomials $r_j$, for some $L_i \neq \emptyset, [r]$. Therefore each term appearing in the sum in equation (7.6) is of the form

$$\prod_{i=1}^{p} \sum_{j_i=0}^{m_i} f_{j_i}^{(i)}(n)p^{j_i n}$$

for polynomials $f_{j_i}^{(i)}$, where $\sum_i m_i = l(T)$ and some $m_i$ can be 0. So the right hand side of equation (7.6) is of the form $\sum_{j=0}^{l(T)} q_j(n)p^{jn}$ for polynomials $q_j$. This means we have the required recurrence relation for $A_T^{(p)}(n)$ and applying Lemma 7.2 gives the result for $A_T^{(p)}(n)$. Finally, we use (7.1) and Lemma 7.2 which gives the result for $C_T^{(p)}(n)$. $\qquad\square$

## 7.2 The leading terms of $A_T^{(p)}(n)$

We now generalise the results of Chapter 6, giving the leading terms of $A_T^{(p)}(n)$, and therefore the asymptotics of $A_T^{(p)}(n)/C_T^{(p)}(n)$. We will require the following generalisation of Lemma 6.1.

**Lemma 7.3.** *The recurrence relation*

$$y_n - py_{n-1} = \sum_{j=0}^{l} f_j(n)p^{jn},$$

*where each $f_j$ is a polynomial, and the leading term of $f_l(n)$ is $\alpha n^d$, has solution*

$$y_n \sim \begin{cases} \dfrac{\alpha}{d+1}n^{d+1}p^n & \text{if } l = 1 \\[3mm] \dfrac{p^{l-1}}{p^{l-1}-1}\alpha n^d p^{ln} & \text{if } l \geq 2. \end{cases} \tag{7.7}$$

*Furthermore, if $d > 0$ and the leading two terms of $f_l(n)$ are $\alpha n^d + \beta n^{d-1}$, then the solution is*

$$y_n \sim \begin{cases} \left(\dfrac{\alpha}{d+1}n^{d+1} + \left(\dfrac{\beta}{d}+\dfrac{\alpha}{2}\right)n^d\right)p^n & \text{if } l = 1 \\[3mm] \dfrac{p^{l-1}}{p^{l-1}-1}\left(\alpha n^d + \left(\beta - \dfrac{d\alpha}{p^{l-1}-1}\right)n^{d-1}\right)p^{ln} & \text{if } l \geq 2. \end{cases} \tag{7.8}$$

$\qquad\square$

**Theorem 7.4.** *The leading polynomial $q_{l(T)}(n)$ in the expression*

$$A_T^{(p)}(n) = \sum_{j=0}^{l(T)} q_j(n)p^{jn}$$

*has degree $d(T)$, where $d(T) = |\{x \in T : x \text{ not the root or a leaf, } D[x] \text{ a chain}\}|$. The coefficient $\alpha_T^{(p)}$ of $n^{d(T)}$ satisfies the following equations.*

*If $T$ is the 2-element chain, then $\alpha_T^{(p)} = 1/(p-1)$. Otherwise, if the root of $T$ has $r$ lower covers, then*

$$\alpha_T^{(p)} = \begin{cases} \dfrac{\alpha_{D[x_1]}^{(p)}}{d(T)} & T \text{ a chain,} \quad r = 1 \\[3ex] \dfrac{\alpha_{D[x_1]}^{(p)}}{p^{l(T)-1} - 1} & T \text{ not a chain,} \quad r = 1 \\[3ex] \dfrac{\displaystyle\sum_{\substack{(L_1,\dots,L_p)\vdash[r] \\ L_i \neq [r]}} \rho \left[ \prod_{i:L_i \neq \emptyset} \alpha_{T_{L_i}}^{(p)} \right]}{p^{l(T)-1} - 1} & r > 1 \end{cases} \tag{7.9}$$

*where $\rho = p^{\sum_{i:|L_i|=1}(l(T_{L_i})-1)-1}$. Moreover, if $d(T) > 0$ the coefficient $\beta_T^{(p)}$ of $n^{d(T)-1}$ satisfies the following equations.*

*If $T$ is the 3-element chain, then $\beta_T^{(p)} = (-2p+1)/(p-1)^2$. Otherwise, if the root of $T$ has $r$ lower covers, then*

$$\beta_T^{(p)} = \begin{cases} \dfrac{\beta_{D[x_1]}^{(p)}}{d(T)-1} - \dfrac{d(T)\alpha_T^{(p)}}{2} & T \text{ a chain,} \quad r = 1 \\[3ex] \dfrac{\beta_{D[x_1]}^{(p)} - d(T)\alpha_T^{(p)}p^{l(T)-1}}{p^{l(T)-1} - 1} & T \text{ not a chain,} \quad r = 1 \\[3ex] \dfrac{\displaystyle\sum_{\substack{(L_1,\dots,L_p)\vdash[r] \\ L_i \neq [r]}} \rho \left[ \sum_{j:L_j \neq \emptyset}\left(\beta_{T_{L_j}}^{(p)} \prod_{\substack{i:L_i \neq \emptyset \\ i \neq j}} \alpha_{T_{L_i}}^{(p)}\right) - \left(\prod_{i:L_i \neq \emptyset} \alpha_{T_{L_i}}^{(p)}\right)\sum_{i:|L_i|>1} d(T_{L_i}) \right]}{p^{l(T)-1} - 1} \\ \quad - \dfrac{d(T)\alpha_T^{(p)}}{p^{l(T)-1} - 1} & r > 1 \end{cases} \tag{7.10}$$

*where $\beta_S = 0$ for any subtree $S \subseteq T$ with $d(S) = 0$.*

**Proof.** Here we prove only that the degree of $q_{l(T)}$ is $d(T)$ and that $\alpha_T^{(p)}$ is as claimed. The proof method is naturally very similar to the method used in the proof of Theorem 6.2, which covers the case $p = 2$. The proof that $\beta_T^{(p)}$ is as claimed can be obtained by considering the coefficient of $n^{d(T)-1}p^{l(T)n}$ in the calculations below, and using (7.8) when applying Lemma 7.3.

We proceed by induction on $|T|$. For $|T| = 2$ we know that $A_T^{(p)}(n) = (p^n - p)/(p-1)$ and for this tree $d(T) = 0$, $l(T) = 1$, so $q_{l(T)}(n) = 1/(p-1)$ a polynomial of degree 0, with leading coefficient equal to $1/(p-1)$. That is, $\alpha_T^{(p)} = 1/(p-1)$ as claimed.

Suppose the result is true for all $T$ with $|T| < k$ and let $T$ be any tree with $|T| = k$. If $T$ has one lower cover, $x_1$, then $A_T^{(p)}(n) = pC_{D[x_1]}^{(p)}(n-1)$. By Theorem 7.1, and our inductive hypothesis, we know that

$$A_{D[x_1]}^{(p)}(n) \sim \alpha_{D[x_1]}^{(p)} n^{d(D[x_1])} p^{l(D[x_1])n}.$$

If $T$ is a chain, then $l(T) = l(D[x_1]) = 1$ and $d(T) = d(D[x_1]) + 1$, so $C_{D[x_1]}^{(p)}(n)$ satisfies the recurrence (7.1), which is of the form in Lemma 7.3 with $\alpha = \alpha_{D[x_1]}^{(p)}$, $d = d(D[x_1])$ and $l = l(D[x_1])$. So, by (7.7)

$$C_{D[x_1]}^{(p)}(n) \sim \frac{\alpha_{D[x_1]}^{(p)}}{d(D[x_1]) + 1} n^{d(D[x_1])+1} p^n = \frac{\alpha_{D[x_1]}^{(p)}}{d(T)} n^{d(T)} p^n.$$

So

$$A_T^{(p)}(n) = pC_{D[x_1]}^{(p)}(n-1) \sim p\frac{\alpha_{D[x_1]}^{(p)}}{d(T)}(n-1)^{d(T)} p^{n-1} = \frac{\alpha_{D[x_1]}^{(p)}}{d(T)}(n-1)^{d(T)} p^n.$$

Therefore $q_{l(T)}$ has degree $d(T)$ and $\alpha_T^{(p)} = \alpha_{D[x_1]}^{(p)}/d(T)$, as claimed. If $T$ is not a chain, then $l(T) = l(D[x_1]) > 1$ and $d(T) = d(D[x_1])$, so again $C_{D[x_1]}^{(p)}(n)$ satisfies a recurrence of the form in Lemma 7.3 with $\alpha = \alpha_{D[x_1]}^{(p)}$, $d = d(T)$ and $l = l(T)$. So, by (7.7),

$$C_{D[x_1]}^{(p)}(n) \sim \frac{p^{l(T)-1}}{p^{l(T)-1} - 1} \alpha_{D[x_1]}^{(p)} n^{d(T)} p^{l(T)n}.$$

So

$$A_T^{(p)}(n) = pC_{D[x_1]}^{(p)}(n-1) \sim p\frac{p^{l(T)-1}}{p^{l(T)-1}-1}\alpha_{D[x_1]}^{(p)}(n-1)^{d(T)}p^{l(T)(n-1)}$$

$$= \frac{\alpha_{D[x_1]}^{(p)}}{p^{l(T)-1}-1}(n-1)^{d(T)}p^{l(T)n}.$$

Therefore $q_{l(T)}$ has degree $d(T)$ and $\alpha_T^{(p)} = \alpha_{D[x_1]}^{(p)}/(p^{l(T)-1}-1)$, as claimed.

If $T$ has $r > 1$ lower covers, then $A_T^{(p)}(n)$ satisfies the recurrence

$$A_T^{(p)}(n) - pA_T^{(p)}(n-1) = \sum_{\substack{L_1,\dots,L_p:\\(L_1,\dots,L_p)\vdash[r],\\L_i\ne[r]}} \prod_{i=1}^{p} A_{T_{L_i}}^{(p)-}(n).$$

Since $A_{T_{L_i}}^{(p)-}(n)$ is equal to 1 when $L_i = \emptyset$, the equation above becomes

$$A_T^{(p)}(n) - pA_T^{(p)}(n-1) = \sum_{\substack{(L_1,\dots,L_p)\vdash[r],\, i:L_i\ne\emptyset\\L_i\ne[r]}} \prod A_{T_{L_i}}^{(p)-}(n).$$

For $L_i \ne \emptyset$, the term $A_{T_{L_i}}^{(p)-}(n)$ is either $A_{T_{L_i}}^{(p)}(n)/p$ or $A_{T_{L_i}}^{(p)}(n-1)$, depending on whether $L_i$ is a singleton or not. By our inductive hypothesis the leading term of $A_{T_{L_i}}^{(p)}(n)$ is $\alpha_{T_{L_i}}^{(p)} n^{d(T_{L_i})}p^{l(T_{L_i})n}$, so we have

$$\prod_{i:L_i\ne\emptyset} A_{T_{L_i}}^{(p)-}(n) = \prod_{i:|L_i|=1} \alpha_{T_{L_i}}^{(p)} n^{d(T_{L_i})}p^{l(T_{L_i})n}/p \prod_{i:|L_i|>1} \alpha_{T_{L_i}}^{(p)}(n-1)^{d(T_{L_i})}p^{l(T_{L_i})(n-1)}$$

$$\sim p^{-\sum_{i:|L_i|=1}1 - \sum_{i:|L_i|>1}l(T_{L_i})} \left[\prod_{i:L_i\ne\emptyset}\alpha_{T_{L_i}}^{(p)}\right] n^{d(T)}p^{l(T)n}$$

$$= \frac{\rho}{p^{l(T)-1}}\left[\prod_{i:L_i\ne\emptyset}\alpha_{T_{L_i}}^{(p)}\right] n^{d(T)}p^{l(T)n}$$

where $\rho = p^{\sum_{i:|L_i|=1}(l(T_{L_i})-1)-1}$. Therefore $A_T^{(p)}$ satisfies a recurrence of the form in Lemma 7.3, with

$$\alpha = \sum_{\substack{(L_1,\dots,L_p)\vdash[r],\\L_i\ne[r]}} \frac{\rho}{p^{l(T)-1}}\left[\prod_{i:L_i\ne\emptyset}\alpha_{T_{L_i}}^{(p)}\right],$$

$d = d(T)$, $l = l(T) > 1$. So by (7.7),

$$A_T^{(p)}(n) \sim \frac{p^{l(T)-1}}{p^{l(T)-1} - 1} \sum_{\substack{(L_1,\ldots,L_p) \vdash [r], \\ L_i \neq [r]}} \frac{\rho}{p^{l(T)-1}} \left[ \prod_{i : L_i \neq \emptyset} \alpha_{T_{L_i}}^{(p)} \right] n^{d(T)} p^{l(T)}$$

$$= \frac{\displaystyle\sum_{\substack{(L_1,\ldots,L_p) \vdash [r], \\ L_i \neq [r]}} \rho \left[ \prod_{i : L_i \neq \emptyset} \alpha_{T_{L_i}}^{(p)} \right]}{p^{l(T)-1} - 1} n^{d(T)} p^{l(T)}.$$

Therefore $q_{l(T)}$ has degree $d(T)$ and $\alpha_T^{(p)}$ is as claimed.      □

Note that we can use equations (7.9) and (7.10) to explicitly calculate $\alpha_T^{(p)}$ and $\beta_T^{(p)}$ for a particular tree $T$, and particular $p$, but the calculations would be extremely cumbersome. Even without expressions for $\alpha_T^{(p)}$ and $\beta_T^{(p)}$ we can see that the dominant term accounts for most embeddings, as in the complete binary tree case.

## 7.3 Typical embeddings $T$ into $T_p^n$

As in section 6.2 we describe a "typical" embedding of $T$ into $T_p^n$, which shows that the leading term given in the previous section gives the lower bound $A_T^{(p)}(n) = \Omega(n^{d(T)} p^{l(T)n})$.

If $T$ is a chain, we can count the embeddings that map the root of $T$ to $1_n$ and the leaf of $T$ to some leaf of $T_p^n$. We have $p^{n-1}$ choices for where to map the leaf. Once we have fixed the leaf of $T_p^n$, this defines a path from $1_n$ to the leaf of $T_p^n$. This gives a choice of $n - 2$ elements of $T_p^n$ into which we can map the $d(T)$ lower bead elements of $T$. So, asymptotically we have $\Theta(n^{d(T)})$ choices for where to map the $d(T)$ lower bead elements. Therefore $A_T^{(p)}(n) = \Omega(n^{d(T)} p^n)$, and since $l(T) = 1$ we have that $A_T(n) = \Omega(n^{d(T)} p^{l(T)n})$ for $T$ a chain.

For $T$ not a chain, so there exist branching elements of $T$, let $\phi$ be some embedding which maps the root of $T$ to $1_n$, and maps the branching elements of $T$ to as

high a level of $T_p^n$ as possible. Consider, for large $n$, the number of embeddings of $T$ into $T_p^n$ that agree with this fixed $\phi$ on the root, branching elements and upper bead elements. Let us only consider those embeddings which map the leaves of $T$ to the leaves of $T_p^n$. If a leaf $y$ is joined to a branching element $x$ by a chain of lower bead elements, then note that $\phi(x)$ is a fixed distance from the root of $T_p^n$, so that $\phi(x)$ is in level $n - k_x$ of $T_p^n$, where $k_x$ is a constant independent of $n$. So, given $\phi$, the leaf $y$ can be mapped to $p^{-k_x}p^{n-1}$ leaves in $T_p^n$. So, the total number of choices for all the leaves is asymptotically $\Theta(p^{l(T)n})$. (The over-counting due to the possibility that two leaves that are below the same branching point are mapped to the same leaf of $T_p^n$ is negligible for large $n$.) It remains to choose where to map the lower bead elements. However, in a similar way to the case where $T$ is a chain, a lower bead element on the chain between the branching point $x$ and the leaf $y$ must be mapped to an element on the path between the images of $x$ and $y$. Since $x$ is mapped to level $n - k_x$, and $y$ to a leaf, the path has $n - k_x - 2$ elements, with $k_x$ independent of $n$. Since there are $d(T)$ lower bead elements, we have asymptotically $\Theta(n^{d(T)})$ choices for where to map the lower bead elements. (Again, this is an over-count due to the possibility that two lower bead elements that are below the same branching element but above different leaves are mapped to the same element of $T_p^n$. However, this is negligible because typically the lower bead elements will not be mapped within $O(1)$ of a branching element.) So, the number of embeddings that agree with $\phi$ is asymptotically $\Omega(n^{d(T)}p^{l(T)n})$, and we have $A_T^{(p)}(n) = \Omega(n^{d(T)}p^{l(T)n})$ for $T$ not a chain.

By Lemma 7.3 we also have the asymptotic behaviour of $C_T^{(p)}(n)$, given in the following corollary.

**Corollary 7.5.** *For any tree $T$ with $l(T) = 1$ the number of embeddings of $T$ into $T_p^n$ is asymptotically*

$$C_T^{(p)}(n) \sim \frac{\alpha_T^{(p)}}{d(T) + 1} n^{d(T)+1} p^n$$

and if $d(T) > 0$ then

$$C_T^{(p)}(n) \sim \left( \frac{\alpha_T^{(p)}}{d(T)+1} n^{d(T)+1} + \left( \frac{\beta_T^{(p)}}{d(T)} + \frac{\alpha_T^{(p)}}{2} \right) n^{d(T)} \right) p^n.$$

For any tree with $l(T) > 1$ the number of embeddings of $T$ into $T_p^n$ is asymptotically

$$C_T^{(p)}(n) \sim \frac{p^{l(T)-1}}{p^{l(T)-1} - 1} \alpha_T^{(p)} n^{d(T)} p^{l(T)n}$$

and if $d(T) > 0$ then

$$C_T^{(p)}(n) \sim \frac{p^{l(T)-1}}{p^{l(T)-1} - 1} \left( \alpha_T^{(p)} n^{d(T)} + \left( \beta_T^{(p)} - \frac{d(T)\alpha_T^{(p)}}{p^{l(T)-1} - 1} \right) n^{d(T)-1} \right) p^{l(T)n}.$$

This tells us that for a tree $T$ not a chain, a typical embedding of $T$ into $T_p^n$ maps the leaves of $T$ to the low levels of $T_p^n$, the branching points and upper bead elements of $T$ to the high levels of $T_p^n$, and the lower bead elements of $T$ will be mapped to elements spread roughly evenly along the paths in $T_p^n$ defined by the images of branching elements and leaves of $T$, as explained earlier. There are $\Theta(n^{d(T)} p^{l(T)n})$ of these embeddings.

For $T$ a chain, a typical embedding maps the leaf of $T$ to a low level of $T_p^n$, and the remaining elements of $T$ are mapped to elements spread roughly evenly on the path from $1_n$ to image of the leaf in $T^n$. Here the root is not necessarily mapped to $1_n$, and the root can be thought of as a lower bead element, so there are $d(T) + 1$ elements to position on this path. So, we get $\Theta(n^{d(T)+1} p^n)$ of these embeddings.

# 7.4   Asymptotics of $A_T^{(p)}(n)/C_T^{(p)}(n)$

We have the following extension to a result of Kubicki, Lehel, and Morayne, which follows immediately from the asymptotic expressions for $A_T^{(p)}(n)$ and $C_T^{(p)}(n)$ given by Theorem 7.4 and Corollary 7.5.

**Proposition 7.6.**

$$\lim_{n \to \infty} \frac{A_T^{(p)}(n)}{C_T^{(p)}(n)} = \frac{p^{l(T)-1} - 1}{p^{l(T)-1}} \qquad \square$$

This provides the following extension to the asymptotic inequality of Kubicki, Lehel and Morayne.

**Corollary 7.7.** *For any $n$ and $p$ and any trees $T_1, T_2$ such that $T_2$ contains a subposet isomorphic to $T_1$, we have*

$$\lim_{n \to \infty} \frac{A_{T_1}^{(p)}(n)}{C_{T_1}^{(p)}(n)} \le \lim_{n \to \infty} \frac{A_{T_2}^{(p)}(n)}{C_{T_2}^{(p)}(n)}. \tag{7.11}$$

**Proof.** This follows immediately from Proposition 7.6, since we have $l(T_1) \le l(T_2)$.

$\square$

We have the following asymptotic behaviour of $A_T^{(p)}(n)/C_T^{(p)}(n)$, similar to the result for the binary complete tree.

**Theorem 7.8.** *For any tree $T$ with $l(T) > 1$ and $d(T) > 0$, we have*

$$\frac{A_T^{(p)}(n)}{C_T^{(p)}(n)} = 1 - \frac{1}{p^{l(T)-1}} \left( 1 - \frac{d(T)}{n} + \frac{\binom{d(T)}{2}}{n^2} + \frac{b_T^{(p)}}{n^2} \right) + o(n^{-2}) \tag{7.12}$$

*where*

$$b_T^{(p)} = \frac{\beta_T^{(p)}}{\alpha_T^{(p)}} - \frac{d(T)}{p^{l(T)-1} - 1}. \tag{7.13}$$

*For any tree $T$ with $l(T) > 1$ and $d(T) = 0$, we have*

$$\frac{A_T^{(p)}(n)}{C_T^{(p)}(n)} = 1 - \frac{1}{p^{l(T)-1}} + o(n^{-1}). \tag{7.14}$$

*For any tree $T$ with $l(T) = 1$, we have*

$$\frac{A_T^{(p)}(n)}{C_T^{(p)}(n)} = \frac{d(T) + 1}{n} + o(n^{-1}). \tag{7.15}$$

$\square$

**Corollary 7.9.** *For any two trees $T_1, T_2$, if either*

*(i) $l(T_1) > l(T_2)$, or*

*(ii) $l(T_1) = l(T_2)$ and $d(T_1) > d(T_2)$, or*

*(iii) $l(T_1) = l(T_2)$, $d(T_1) = d(T_2) > 0$ and $\alpha_{T_1}^{(p)}/\beta_{T_1}^{(p)} > \alpha_{T_2}^{(p)}/\beta_{T_2}^{(p)}$,*

*then there exists an integer $n_0$ such that*

$$\frac{A_{T_1}^{(p)}(n)}{C_{T_1}^{(p)}(n)} > \frac{A_{T_2}^{(p)}(n)}{C_{T_2}^{(p)}(n)}$$

*for all $n \geq n_0$.*        □

This implies that to find examples of trees $T_1, T_2$ with $T_2$ containing a subposet isomorphic to $T_1$ and satisfying

$$\frac{A_{T_1}^{(p)}(n)}{C_{T_1}^{(p)}(n)} > \frac{A_{T_2}^{(p)}(n)}{C_{T_2}^{(p)}(n)}$$

for all $n$ greater than some $n_0$, we need only consider pairs with $l(T_1) = l(T_2)$, $d(T_1) = d(T_2) > 0$ and $\alpha_{T_1}^{(p)}/\beta_{T_1}^{(p)} > \alpha_{T_2}^{(p)}/\beta_{T_2}^{(p)}$.

We finish this chapter by showing that the pair of trees $T_1, T_2$ in Figure 5.1 on page 106 are an example of such a pair. The following calculations are very similar to those in the proof of Theorem 6.6, but using the more complicated expressions (7.9) and (7.10) for $\alpha_T^{(p)}$ and $\beta_T^{(p)}$. Indeed, it would be possible to generalise Theorem 6.6 to give a whole family of examples, however the calculations would be rather more involved.

**Theorem 7.10.** *Let $T_1, T_2$ be trees as depicted in Figure 5.1. There exists some $n_0$ such that*

$$\frac{A_{T_1}^{(p)}(n)}{C_{T_1}^{(p)}(n)} > \frac{A_{T_2}^{(p)}(n)}{C_{T_2}^{(p)}(n)}$$

*for all $n \geq n_0$.*

**Proof.** Note that $l(T_1) = l(T_2) = 3$ and $d(T_1) = d(T_2) = 1$. So, by Corollary 7.9, it is enough to show that

$$\frac{\alpha_{T_1}^{(p)}}{\beta_{T_1}^{(p)}} > \frac{\alpha_{T_2}^{(p)}}{\beta_{T_2}^{(p)}}.$$

As in the proof of Theorem 6.6, we will calculate $\alpha_{T_1}^{(p)}\beta_{T_2}^{(p)} - \alpha_{T_2}^{(p)}\beta_{T_1}^{(p)}$ and show that it is positive.

Let the lower covers of the root of $T_1$ be $x_1, x_2, x_3$ and let $y$ be the extra element of $T_2$, so that $T_1 = T_2 \setminus \{y\}$. For each tree $T = T_1, T_2$, we write $\alpha_T^{(p)}$ and $\beta_T^{(p)}$ in

$T_{\{1\}}$     $T_{\{2\}} = T_{\{3\}}$     $T_{\{1,2\}} = T_{\{1,3\}}$     $T_{\{2,3\}}$     $T_{\{y\}}$
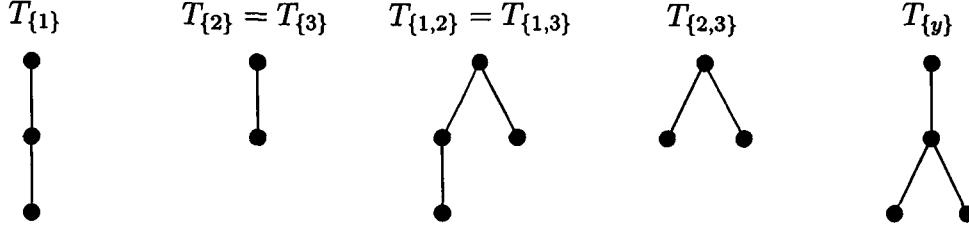
Figure 7.1: Subtrees of $T_1$ and $T_2$

terms of $\alpha_S^{(p)}$ and $\beta_S^{(p)}$ for the subtrees $S$ defined in Figure 7.1. For ease of notation, we write $\alpha_1^{(p)}$ for $\alpha_{T_{\{1\}}}^{(p)}$, $\alpha_{12}^{(p)}$ for $\alpha_{T_{\{1,2\}}}^{(p)}$, etc., and similarly for $\beta_S^{(p)}$.

Also, note that $d(T_{\{y\}}) = d(T_{\{2,3\}}) = d(T_{\{2\}}) = d(T_{\{3\}}) = 0$ and therefore $\beta_y^{(p)} = \beta_{23}^{(p)} = \beta_2^{(p)} = \beta_3^{(p)} = 0$, which will simplify the calculations.

Using (7.10), we have

$$\beta_{T_1}^{(p)} = \frac{p-1}{p^2-1}\left[\begin{array}{c} \beta_1^{(p)}\alpha_{23}^{(p)} + \beta_{13}^{(p)}\alpha_2^{(p)} - \alpha_2^{(p)}\alpha_{13}^{(p)} + \beta_{12}^{(p)}\alpha_3^{(p)} - \alpha_3^{(p)}\alpha_{12}^{(p)} \\ + (p-2)\beta_1^{(p)}\alpha_2^{(p)}\alpha_3^{(p)} \end{array}\right] - \frac{\alpha_{T_1}^{(p)}}{p^2-1},$$

$$\beta_{T_2}^{(p)} = \frac{p(p-1)\beta_1^{(p)}\alpha_y^{(p)}}{p^2-1} - \frac{\alpha_{T_2}^{(p)}}{p^2-1}$$

so

$$\alpha_{T_1}^{(p)}\beta_{T_2}^{(p)} - \alpha_{T_2}^{(p)}\beta_{T_1}^{(p)} = \frac{p-1}{p^2-1}\left[\begin{array}{c} p\beta_1^{(p)}\alpha_y^{(p)}\alpha_{T_1}^{(p)} - \left(\beta_1^{(p)}\alpha_{23}^{(p)} + \beta_{13}^{(p)}\alpha_2^{(p)} - \alpha_2^{(p)}\alpha_{13}^{(p)}\right)\alpha_{T_2}^{(p)} \\ - \left(\beta_{12}^{(p)}\alpha_3^{(p)} - \alpha_3^{(p)}\alpha_{12}^{(p)} + (p-2)\beta_1^{(p)}\alpha_2^{(p)}\alpha_3^{(p)}\right)\alpha_{T_2}^{(p)} \end{array}\right]$$

and using (7.9), we have

$$\alpha_{T_1}^{(p)} = \frac{p-1}{p^2-1}\left[\alpha_1^{(p)}\alpha_{23}^{(p)} + \alpha_2^{(p)}\alpha_{13}^{(p)} + \alpha_3^{(p)}\alpha_{12}^{(p)} + (p-2)\alpha_1^{(p)}\alpha_2^{(p)}\alpha_3^{(p)}\right],$$

$$\alpha_{T_2}^{(p)} = \frac{p(p-1)\alpha_1^{(p)}\alpha_y^{(p)}}{p^2-1}.$$

This gives,

$$
\frac{(\alpha_{T_1}^{(p)}\beta_{T_2}^{(p)} - \alpha_{T_2}^{(p)}\beta_{T_1}^{(p)})(p^2-1)^2}{\alpha_y^{(p)}p(p-1)^2} = \beta_1^{(p)}\left(\alpha_1^{(p)}\alpha_{23}^{(p)} + \alpha_2^{(p)}\alpha_{13}^{(p)} + \alpha_3^{(p)}\alpha_{12}^{(p)}\right)
$$

$$
+ \beta_1^{(p)}(p-2)\alpha_1^{(p)}\alpha_2^{(p)}\alpha_3^{(p)}
$$

$$
- \alpha_1^{(p)}\left(\beta_1^{(p)}\alpha_{23}^{(p)} + \beta_{13}^{(p)}\alpha_2^{(p)} - \alpha_2^{(p)}\alpha_{13}^{(p)}\right)
$$

$$
- \alpha_1^{(p)}\left(\beta_{12}^{(p)}\alpha_3^{(p)} - \alpha_3^{(p)}\alpha_{12}^{(p)} + (p-2)\beta_1^{(p)}\alpha_2^{(p)}\alpha_3^{(p)}\right)
$$

$$
= \alpha_2^{(p)}\left(\beta_1^{(p)}\alpha_{13}^{(p)} - \alpha_1^{(p)}\beta_{13}^{(p)} + \alpha_1^{(p)}\alpha_{13}^{(p)}\right)
$$

$$
+ \alpha_3^{(p)}\left(\beta_1^{(p)}\alpha_{12}^{(p)} - \alpha_1^{(p)}\beta_{12}^{(p)} + \alpha_1^{(p)}\alpha_{12}^{(p)}\right).
$$

Finally, we have

$$
\beta_{13}^{(p)} = \frac{(p-1)\beta_1^{(p)}\alpha_3^{(p)} - \alpha_{13}^{(p)}}{p-1} \quad \text{and} \quad \alpha_{13}^{(p)} = \alpha_1^{(p)}\alpha_3^{(p)}
$$

so

$$
\beta_1^{(p)}\alpha_{13}^{(p)} - \alpha_1^{(p)}\beta_{13}^{(p)} + \alpha_1^{(p)}\alpha_{13}^{(p)} = \beta_1^{(p)}\alpha_1^{(p)}\alpha_3^{(p)} - \alpha_1^{(p)}\frac{(p-1)\beta_1^{(p)}\alpha_3^{(p)} - \alpha_{13}^{(p)}}{p-1} + \alpha_1^{(p)}\alpha_{13}^{(p)}
$$

$$
= \frac{p}{p-1}\alpha_1^{(p)}\alpha_{13}^{(p)}
$$

and similarly

$$
\beta_1^{(p)}\alpha_{12}^{(p)} - \alpha_1^{(p)}\beta_{12}^{(p)} + \alpha_1^{(p)}\alpha_{12}^{(p)} = \frac{p}{p-1}\alpha_1^{(p)}\alpha_{12}^{(p)}.
$$

Therefore

$$
\alpha_{T_1}^{(p)}\beta_{T_2}^{(p)} - \alpha_{T_2}^{(p)}\beta_{T_1}^{(p)} = \frac{\alpha_y^{(p)}p(p-1)^2}{(p^2-1)^2}\frac{p}{p-1}(\alpha_1^{(p)}\alpha_2^{(p)}\alpha_{13}^{(p)} + \alpha_1^{(p)}\alpha_3^{(p)}\alpha_{12}^{(p)})
$$

$$
= \frac{\alpha_y^{(p)}\alpha_1^{(p)}p^2(p-1)}{(p^2-1)^2}(\alpha_2^{(p)}\alpha_{13}^{(p)} + \alpha_3^{(p)}\alpha_{12}^{(p)})
$$

$$
= \frac{p}{p^2-1}\alpha_{T_2}^{(p)}(\alpha_2^{(p)}\alpha_{13}^{(p)} + \alpha_3^{(p)}\alpha_{12}^{(p)})
$$

which is positive, since $\alpha_T^{(p)}$ is positive for all $T$. $\qquad\square$

Note that the above expression for $\alpha_{T_1}^{(p)}\beta_{T_2}^{(p)} - \alpha_{T_2}^{(p)}\beta_{T_1}^{(p)}$, with $p = 2$, corresponds exactly with the expression for $\alpha_{T_1}\beta_{T_2} - \alpha_{T_2}\beta_{T_1}$ found at the end of the proof of Theorem 6.6, for the specific trees $T = T_1$ and $T' = T_2$. In fact, the calculations throughout the two proofs are very similar; the main difference is that for $p > 2$, the set [3] can be partitioned into three sets $\{1\} \cup \{2\} \cup \{3\}$, which is not possible if

$p = 2$. These partitions account for the terms with the prefactor of $(p - 2)$ (which vanish for $p = 2$), and we note that these terms apparently cancel when calculating

$$\alpha_{T_1}^{(p)}\beta_{T_2}^{(p)} - \alpha_{T_2}^{(p)}\beta_{T_1}^{(p)}.$$

As mentioned earlier, this hints at a possible generalisation to Theorem 6.6 which would give a family of pairs of trees $T \subseteq T'$ with

$$\frac{A_T^{(p)}(n)}{C_T^{(p)}(n)} > \frac{A_{T'}^{(p)}(n)}{C_{T'}^{(p)}(n)}$$

for sufficiently large $n$. However, more important than finding many of these pairs is the fact that there is at least one such pair of trees that are ternary. Theorem 7.10 tells us that if we want to generalise Theorem 4.1 of Kubicki, Lehel, and Morayne, to embeddings of trees into the complete $p$-ary tree, we must keep the condition on the trees being binary. For example, we do not have, as might at first be hoped, that the result is true for embeddings of $p$-ary trees into the complete $p$-ary tree (nor even for embeddings of ternary trees into the complete $p$-ary tree). In the next chapter we show that the result *is* true for embeddings of binary trees into the complete $p$-ary tree.

# Chapter 8

# Generalisations of Theorem 4.1

We have shown that Theorem 4.1, of Kubicki, Lehel and Morayne, stating that

$$\frac{A_{T_1}(n)}{C_{T_1}(n)} \leq \frac{A_{T_2}(n)}{C_{T_2}(n)}$$

for binary trees $T_1, T_2$ such that $T_2$ contains a subposet isomorphic to $T_1$, does not extend to arbitrary trees $T_1 \subseteq T_2$. Here, we look at generalisations of the result in other directions, for example by looking at embeddings of binary trees into the complete $p$-ary tree, for any $p \geq 2$. We will also generalise the result to order-preserving maps of *arbitrary* trees into the complete $p$-ary tree.

As explained in the previous chapter, we cannot generalise Theorem 4.1 to embeddings of arbitrary trees into the complete $p$-ary tree. In this regard, we have the best possible result, that Theorem 4.1 generalises to embeddings of binary trees into the complete $p$-ary tree.

## 8.1 Embeddings of binary trees into the complete $p$-ary tree

Recall that $T_p^n$ is the complete $p$-ary tree of height $n$, with root $1_n$, and we write

$A_T^{(p)}(n)$ for the number of embeddings of $T$ into $T_p^n$ that map the root $1_T$ of $T$ to $1_n$ and $C_T^{(p)}(n)$ for the total number of embeddings of $T$ into $T_p^n$. We prove the result that

$$\frac{A_{T_1}^{(p)}(n)}{C_{T_1}^{(p)}(n)} \leq \frac{A_{T_2}^{(p)}(n)}{C_{T_2}^{(p)}(n)}$$

for binary trees $T_1, T_2$ such that $T_2$ contains a subposet isomorphic to $T_1$. We do so by defining an appropriate distributive lattice and then applying the FKG-inequality. The FKG-inequality is a powerful corollary of the Four Functions Theorem by Ahlswede and Daykin. See, for example, [3] for a background to the FKG-inequality and examples of its use in probabilistic combinatorics. We state a form of the inequality that we will use repeatedly.

**Theorem 8.1** (Fortuin, Kasteleyn, Ginibre (1971)). *If $(\mathcal{F}, <)$ is a finite distributive lattice and if $\alpha, \beta$ are both increasing (or both decreasing) non-negative functions on $\mathcal{F}$ and $\mu$ is a non-negative function on $\mathcal{F}$ such that $\mu(f)\mu(g) \leq \mu(f \vee g)\mu(f \wedge g)$ for all $f, g \in \mathcal{F}$, then*

$$\sum_{f \in \mathcal{F}} \mu(f)\alpha(f) \sum_{f \in \mathcal{F}} \mu(f)\beta(f) \leq \sum_{f \in \mathcal{F}} \mu(f) \sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\beta(f) \tag{8.1}$$

A function $\mu$ on a lattice $\mathcal{F}$ is said to be *log-supermodular* if

$$\mu(f)\mu(g) \leq \mu(f \vee g)\mu(f \wedge g) \quad \text{for all } f, g \in \mathcal{F}. \tag{8.2}$$

The power of this result means the inequality $A_{T_1}^{(p)}(n)/C_{T_1}^{(p)}(n) \leq A_{T_2}^{(p)}(n)/C_{T_2}^{(p)}(n)$ can be viewed as just one of many correlation inequalities for embeddings of binary trees into complete trees. We define an appropriate distributive lattice $\mathcal{F}$ and log-supermodular function $\mu$ so that $\sum_{f \in \mathcal{F}} \mu(f)$ equals the number of embeddings into $T_p^n$. Then we have the FKG-inequality (8.1) for *any* pair of increasing functions $\alpha, \beta$. As we will see, the definition of the lattice $\mathcal{F}$ means that the indicator functions of events like "the root of $T$ is mapped to $1_n$" or "element $x \in T$ is mapped to a high level of $T_p^n$" will be increasing on $\mathcal{F}$. The FKG-inequality then tells us that events like this are positively correlated, i.e., the probability that one event occurs increases if we condition on the other event occurring.

We only need consider the case where $T_1$ and $T_2$ differ by one element, since we can reduce to this case by the following lemmas. Lemma 8.2 is obvious, and the proof of Lemma 8.3 can be found in [17].

**Lemma 8.2.** *Given a binary tree, the following types of operation produce another binary tree with one element fewer.*

*(a) Removing a leaf,*

*(b) Removing the lower cover of an element that has exactly one lower cover.*    □

Note that if an element has exactly one lower cover and the lower cover is also a leaf, removing this leaf can be considered as an operation of both types. Also, note that we can think of operation (b) as contracting the edge between the element and its lower cover, that is, identifying them in the new tree.

**Lemma 8.3.** *If $T_1$ and $T_2$ are binary trees and $T_2$ contains a subposet isomorphic to $T_1$, then there is a sequence of operations of type (a) and (b) leading from $T_2$ to an isomorphic copy of $T_1$ through binary trees.*    □

**Theorem 8.4.** *If $T_1$ and $T_2$ are binary trees such that $T_2$ contains a subposet isomorphic to $T_1$, then*

$$\frac{A_{T_1}^{(p)}(n)}{C_{T_1}^{(p)}(n)} \leq \frac{A_{T_2}^{(p)}(n)}{C_{T_2}^{(p)}(n)} \tag{8.3}$$

**Proof.** From Lemma 8.3 it is enough to show (8.3) for the particular cases where $T_1$ is isomorphic to the subposet produced from $T_2$ by exactly one operation of either type (a) or (b). Let $m$ be the element removed from $T_2$, and for ease of notation we identify $T_1$ with the subposet $T_2 \setminus \{m\}$.

Firstly, we define a distributive lattice. Write $[n]$ for the chain on the $n$-element set $\{1, 2, \ldots, n\}$ with the natural ordering. For any binary tree $T$, write $\mathcal{F}_T = \mathcal{F}(n; T)$ for the lattice of strict order-preserving maps from $T$ to $[n]$. So $f \in \mathcal{F}_T$ is a function from $T$ to $[n]$ such that $x > y$ in $T$ implies $f(x) > f(y)$ in $[n]$. The

ordering on $\mathcal{F}_T$ is $f \geq g$ if and only if $f(x) \geq g(x)$ for all $x \in T$. The join, $f \vee g$, is the pointwise maximum of $f$ and $g$, and the meet, $f \wedge g$, is the pointwise minimum of $f$ and $g$. It is relatively simple to check that $\mathcal{F}_T$ is a distributive lattice. The easiest way to see this is to note that it is a sublattice of the distributive lattice $[n]^{|T|}$.

We call a function in $\mathcal{F}_T$ a *level function*. If we have an embedding $\phi$ of $T$ into $T_p^n$, we can construct a function $f$ by setting $f(x)$ equal to the level of $\phi(x)$ in $T_p^n$. Since $\phi$ is an embedding, $x > y$ in $T$ implies that the level of $\phi(x)$ is greater than the level of $\phi(y)$, and so $f(x) > f(y)$. Therefore, $f$ is a level function and we say that $\phi$ *corresponds to* $f$. In fact, we can do this for any strict order-preserving map $\phi$ from $T$ to $T_p^n$. For each level function $f \in \mathcal{F}_T$ we can count the number of embeddings from $T$ to $T_p^n$ that correspond to $f$. This defines a function $\mu$ from $\mathcal{F}_T$ to $\mathbb{R}_+$: $\mu(f) = \mu_1(f)\mu_2(f)$ where $\mu_1, \mu_2$ are defined as

$$\mu_1(f) = p^{n-f(1_T)} \prod_{\substack{x>y,\ \text{an edge in } T}} p^{f(x)-f(y)},$$

$$\mu_2(f) = \prod_{\substack{y \in T, \\ y \text{ has 2 lower} \\ \text{covers, } z_1, z_2}} \left(1 - p^{\max\{f(z_1), f(z_2)\}-f(y)}\right).$$

Here, $\mu_1(f)$ counts the number of strict order-preserving maps from $T$ to $T_p^n$ that correspond to the level function $f$. However, a strict order-preserving map from $T$ to $T_p^n$ need not be an embedding of $T$ into $T_p^n$. The term $\mu_2(f)$ is exactly the fraction of those strict order-preserving maps from $T$ to $T_p^n$ corresponding to the level function $f$ that are also embeddings of $T$ into $T_p^n$. To see that $\mu_1(f)$ and $\mu_2(f)$ are as claimed, suppose we are constructing a strict order-preserving map $\phi$ that corresponds to $f$, by choosing the element $\phi(x)$ from level $f(x)$, for each $x$ from the root, $1_T$, downwards. We have $p^{n-f(1_T)}$ choices for $\phi(1_T)$, and then for each edge $x > y$ in $T$, once we have chosen $\phi(x)$ we have $p^{f(x)-f(y)}$ choices for $\phi(y)$. This gives a total of $\mu_1(f)$ strict order-preserving maps. Since we have $\phi(x) > \phi(y)$ for all $x > y$ in $T$, the map $\phi$ is an embedding if $\phi(z_1)$ and $\phi(z_2)$ are incomparable for all elements $z_1$, $z_2$ with a common upper cover in $T$. Let $y$ be some element

of $T$ which has two lower covers $z_1, z_2$ and, without loss of generality, suppose that $f(z_1) \geq f(z_2)$. When constructing $\phi$, once we have chosen $\phi(y)$ and $\phi(z_2)$ (elements in the levels $f(y)$ and $f(z_2)$ respectively), there are $p^{f(y)-f(z_1)}$ choices for $\phi(z_1)$. One of these choices (the element on the path between $\phi(y)$ and $\phi(z_2)$) will give $\phi(z_1) > \phi(z_2)$ in $T_p^n$, meaning that $\phi$ is not an embedding. The other choices mean $\phi(z_1)$ and $\phi(z_2)$ are incomparable as required for $\phi$ to be an embedding. Because of the regularity of $T_p^n$, these numbers are independent of the choice of $\phi(z_2)$, so the fraction of choices which allow $\phi$ to be an embedding is $1 - p^{-(f(y)-f(z_1))}$. So, taking the product over all such $y$ gives the expression $\mu_2(f)$ as the fraction of strict order-preserving maps (corresponding to $f$) that are also embeddings.

**Claim 8.1.** $\mu$ *is log-supermodular on* $\mathcal{F}_T$.

**Proof of Claim 8.1.** Since

$$(f \wedge g)(x) + (f \vee g)(x) = \min(f(x), g(x)) + \max(f(x), g(x)) = f(x) + g(x)$$

for all $x \in T$, we have that $\mu_1(f)\mu_1(g) = \mu_1(f \wedge g)\mu_1(f \vee g)$. So, it is enough to prove (8.2) for $\mu_2$. For each $y \in T$ with two lower covers, $z_1, z_2$, write $\sigma(f) = \max(f(z_1), f(z_2)) - f(y)$. Since $\mu_2$ is a product of terms indexed by such $y$, it is sufficient to prove that

$$(1 - p^{\sigma(f)})(1 - p^{\sigma(g)}) \leq (1 - p^{\sigma(f \wedge g)})(1 - p^{\sigma(f \vee g)}) \tag{8.4}$$

for all $y \in T$ with two lower covers.

Without loss of generality, we can assume that $f(z_1) \geq f(z_2), g(z_1), g(z_2)$. So

$$\sigma(f \wedge g) = \max\{\min(f(z_1), g(z_1)), \min(f(z_2), g(z_2))\} - \min\{f(y), g(y)\}$$
$$= \max\{g(z_1), \min(f(z_2), g(z_2))\} - \min\{f(y), g(y)\}$$

and

$$\sigma(f \vee g) = \max\{\max(f(z_1), g(z_1)), \max(f(z_2), g(z_2))\} - \max\{f(y), g(y)\}$$
$$= f(z_1) - \max\{f(y), g(y)\}$$

which gives

$$\sigma(f \wedge g) + \sigma(f \vee g) = \max\{g(z_1), \min(f(z_2), g(z_2))\} + f(z_1) - f(y) - g(y)$$

$$\leq \max\{g(z_1), g(z_2)\} + f(z_1) - f(y) - g(y)$$

$$= \sigma(f) + \sigma(g)$$

(with equality unless both $g(z_1) < g(z_2)$ and $f(z_2) < g(z_2)$). Moreover, since $\sigma(f \vee g) = f(z_1) - \max\{f(y), g(y)\}$, if $f(y) \geq g(y)$ then $\sigma(f \vee g) = \sigma(f)$ and so $\sigma(f \wedge g) \leq \sigma(g)$ and then (8.4) follows. Otherwise, $f(y) < g(y)$. Set $s = g(y) - f(y) > 0$. Then $\sigma(f \vee g) = f(z_1) - g(y) = \sigma(f) - s$ and $\sigma(f \wedge g) \leq \sigma(g) + s$. Also, $\sigma(g) + s = \max\{g(z_1), g(z_2)\} - g(y) + s \leq f(z_1) - f(y) = \sigma(f)$. So,

$$(1 - p^{\sigma(f \wedge g)})(1 - p^{\sigma(f \vee g)}) \geq (1 - p^{\sigma(g) + s})(1 - p^{\sigma(f) - s})$$

$$= 1 - p^{\sigma(g) + s} - p^{\sigma(f) - s} + p^{\sigma(f) + \sigma(g)}$$

$$\geq 1 - p^{\sigma(g)} - p^{\sigma(f)} + p^{\sigma(f) + \sigma(g)}$$

$$= (1 - p^{\sigma(f)})(1 - p^{\sigma(g)}),$$

where the second inequality holds since the function $\chi : x \mapsto p^x$ is convex for all $x \in \mathbb{R}$, and $\sigma(g) \leq \sigma(g) + s, \sigma(f) - s \leq \sigma(f)$ with $s > 0$.    $\square$

So, we have that $\mu$ is log-supermodular on $\mathcal{F}_T$, and therefore the restriction $\mu'$ of $\mu$ to any sublattice $\mathcal{F}'$ of $\mathcal{F}_T$ is log-supermodular on $\mathcal{F}'$.

We have that the number of embeddings of $T$ into $T_p^n$ is $\sum_{f \in \mathcal{F}_T} \mu(f)$. Also, we can split a tree $T$ at any point and perform similar sums on the two subtrees, as follows. Recall that for $x \in T$, the set $D[x]$ is the down-set of $x$ in $T$. Write $D(x)$ for the set $D[x] \setminus \{x\}$ of elements below $x$ in $T$. Let $x$ be an element of $T$ and define subtrees $S_1 = T \setminus D(x)$ and $S_2 = D[x]$ and consider two lattices $\mathcal{F}_1(k) = \{f \in \mathcal{F}(n; S_1) : f(x) = k\}$ and $\mathcal{F}_2(k) = \{f \in \mathcal{F}(k; S_2) : f(x) = k\}$, where $1 \leq k \leq n$. Then $\sum_{f \in \mathcal{F}_1(k)} \mu(f)$ is the number of embeddings of $S_1$ into $T_p^n$ that map $x$ to an element of $T_p^n$ in level $k$, and $\sum_{f \in \mathcal{F}_2(k)} \mu(f)$ is the number of embeddings of $S_2$ into $T_p^k$ that map $x$ to the root (the only element in level $k$ of $T_p^k$). Consider any pair of embeddings $(\phi_1, \phi_2)$ where $\phi_1$ is an embedding of $S_1$ into $T_p^n$ that maps $x$ to
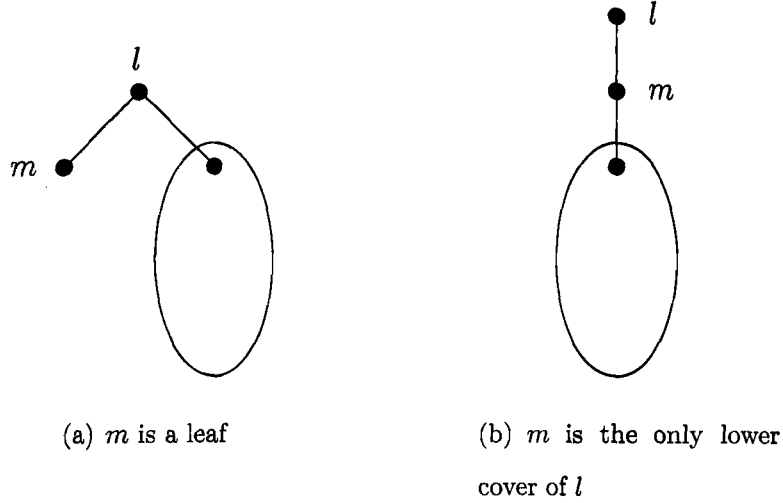
an element in level $k$, and $\phi_2$ is an embedding of $S_2$ into $T_p^k$ that maps $x$ to the root of $T_p^k$. We can construct an embedding $\phi$ of $T$ into $T_p^n$ as follows. For any point $y \in S_1$, define $\phi(y)$ to be $\phi_1(y)$. So, the point $x \in S_1$ is mapped to $\phi(x) = \phi_1(x)$, an element in level $k$. So, $\phi_1$ specifies a unique copy of $T_p^k$ in $T_p^n$, namely the down-set of $\phi_1(x)$ in $T_p^n$. So, for elements $y \in S_2$ define $\phi(y)$ to be the element corresponding to $\phi_2(y)$ in this copy of $T_p^k$. Since the only element in $S_1 \cap S_2$ is $x$ and $\phi_2(x)$ is by definition the root of $T_p^k$, we have a well defined function $\phi$ and this is certainly an embedding of $T$ into $T_p^n$ that maps $x$ to an element in level $k$. Since any embedding of $T$ into $T_p^n$ that maps $x$ to an element in level $k$ can be split into two embeddings by reversing this process, we have that the number of embeddings of $T$ into $T_p^n$ that map $x$ to an element in level $k$ is $\sum_{f \in \mathcal{F}_1(k)} \mu(f) \sum_{g \in \mathcal{F}_2(k)} \mu(g)$ and therefore the total number of embeddings of $T$ into $T^n$ is

$$\sum_{k=1}^{n} \sum_{f \in \mathcal{F}_1(k)} \mu(f) \sum_{g \in \mathcal{F}_2(k)} \mu(g). \tag{8.5}$$

Note that this holds for any element $x$ in $T$.

Recall that $m$ is the point removed from $T_2$ to obtain $T_1$. Let $l$ be the upper cover of $m$ in $T_2$. Write $T_t$ for the subtree $T_1 \setminus D(l)$, and $T_b$ for $D[l]$ as a subtree of $T_1$. Note that we have split $T_1$ into two trees $T_t$ and $T_b$ as explained earlier. Write $T_b+$ for the tree $D[l]$ as a subtree of $T_2$, so that $T_b+ = T_b \cup \{m\}$. Therefore, we have split $T_2$ into two trees $T_t$ and $T_b+$. So, $T_t$ is common to both trees $T_1, T_2$ and $T_b$ and $T_b+$ differ by only one element. Furthermore, since we have that $T_1$ is obtained from $T_2$ either by (a) removing a leaf, or (b) removing the lower cover of an element with exactly one lower cover, we know that either (a) $T_b+$ has the extra element $m$ as a leaf, directly below the root $l$ of $T_b+$, or (b) $T_b+$ has the extra element $m$ as the only lower cover of $l$. (See Figure 8.1.)

Let us look at the sublattice $\mathcal{F}'$ of $\mathcal{F}(n; T_t)$ defined by $\mathcal{F}' = \{f \in \mathcal{F}(n; T_t) : f(l) = k \text{ or } f(l) = k + 1\}$, for $1 \le k < n$. We have $\mu$ defined on $\mathcal{F}'$ as described earlier, and $\mu$ is log-supermodular. Define $\alpha(f) = I\{f(1_{T_t}) = n\}$ as the indicator function of the event $f(1_{T_t}) = n$ and define $\beta(f) = I\{f(l) = k + 1\}$ as the indicator

(a) $m$ is a leaf                    (b) $m$ is the only lower

cover of $l$

Figure 8.1: The two cases for $T_b+$

of the event $f(l) = k + 1$. Both $\alpha$ and $\beta$ are increasing functions, since the sets $\{f : f(1_{T_t}) = n\}$ and $\{f : f(l) = k + 1\}$ are both up-sets of $\mathcal{F}'$.

For $k = 1, \ldots, n$, let $a_k$ be the number of embeddings of $T_t$ into $T_p^n$ that map $l$ to an element in level $k$, and let $b_k$ be the number of embeddings of $T_t$ into $T_p^n$ that map $l$ to an element in level $k$ and map the root $1_{T_t}$ to the root $1_n$. Then,

$$\sum_{f \in \mathcal{F}'} \mu(f)\alpha(f) = b_k + b_{k+1}, \qquad \sum_{f \in \mathcal{F}'} \mu(f) = a_k + a_{k+1},$$

$$\sum_{f \in \mathcal{F}'} \mu(f)\beta(f) = a_{k+1}, \qquad \sum_{f \in \mathcal{F}'} \mu(f)\alpha(f)\beta(f) = b_{k+1},$$

and applying Theorem 8.1 to $\mathcal{F}', \mu, \alpha, \beta$ gives $(b_k + b_{k+1})a_{k+1} \le (a_k + a_{k+1})b_{k+1}$ or

$$\frac{b_k}{a_k} \le \frac{b_{k+1}}{a_{k+1}}$$

for all $k$, $1 \le k < n$.

Now let us look at the trees $T_b$ and $T_b+$. Let $c_k$ be the number of embeddings of $T_b$ into $T_p^k$ that map $l$ to $1_k$, and let $d_k$ be the number of embeddings of $T_b+$ into $T_p^k$ that map $l$ to $1_k$, for $k = 1, \ldots, n$. First consider case (a), where $m$ is a leaf of $T_b+$.

Each embedding of $T_b+$ with $l$ mapped to $1_k$ can be thought of as an extension of an embedding of $T_b$ with $l$ mapped to $1_k$. To extend an embedding of $T_b$ with

$l$ mapped to $1_k$ to an embedding of $T_b+$ with $l$ mapped to $1_k$ we just need to decide where in $T_p^k$ to map $m$, being careful to make sure that the chosen element is incomparable with the image of the other lower cover of $l$. When this other lower cover is mapped to level $k-1$ of $T_p^k$, there are $p^{k-1}-1$ choices for $m$, and if it is mapped to any lower level, then there are more choices for $m$. Since the total number of elements below $1_k$ is equal to $(p^k-p)/(p-1)$, we have that every embedding of $T_b$ with $l$ mapped to $1_k$ can be extended to at least $p^{k-1}-1$ distinct embeddings of $T_b+$ with $l$ mapped to $1_k$, but to at most $(p^k-p)/(p-1)$ distinct embeddings of $T_b+$ with $l$ mapped to $1_k$. Since $p \geq 2$, we have

$$\frac{d_k}{c_k} \leq \frac{p^k-p}{p-1} < p^k - 1 \leq \frac{d_{k+1}}{c_{k+1}}.$$

We now show that $d_k/c_k \leq d_{k+1}/c_{k+1}$ also holds in case (b), again using Theorem 8.1. Let $\mathcal{F}''$ be the sublattice of $\mathcal{F}(k+1; T_b)$ defined as $\mathcal{F}'' = \{f \in \mathcal{F}(k+1; T_b) : f(l) = k$ or $f(l) = k+1\}$, for $1 \leq k < n$. Take $\mu$ defined on this sublattice as before, so that $\mu$ is log-supermodular. Define $\alpha(f) = I\{f(l) = k+1\}$ and define $\beta(f) = (p^{f_{min}} - p)/(p-1)$, where $f_{min} = \min_{x \in T_b} f(x)$. We have that $\alpha$ is increasing on $\mathcal{F}''$, and $f_{min}$ is increasing on $\mathcal{F}''$ therefore $\beta$ is also increasing on $\mathcal{F}''$. Before applying Theorem 8.1 we show what each of the terms in (8.1) is.

There are $p$ elements in level $k$ of $T_p^{k+1}$ whose down-set is a copy of $T_p^k$, so each of the $c_k$ embeddings of $T_b$ into $T_p^k$ that map $l$ to $1_k$ corresponds to $p$ distinct embeddings of $T_b$ into $T_p^{k+1}$ that map $l$ to an element in level $k$. Therefore the sum $\sum_{f \in \mathcal{F}''} \mu(f)$, which counts embeddings of $T_b$ into $T_p^{k+1}$ that map $l$ to an element in level $k$ or $k+1$, equals $pc_k + c_{k+1}$. The sum $\sum_{f \in \mathcal{F}''} \mu(f)\alpha(f)$ equals $c_{k+1}$. The sum $\sum_{f \in \mathcal{F}''} \mu(f)\beta(f)$ counts the number of embeddings of $T_b+$ into $T_p^{k+1}$ that map $l$ to an element in level $k$ or $k+1$. To see this, fix $f$ in $\mathcal{F}''$ and let $\phi$ be an embedding of $T_b$ into $T_p^{k+1}$ that corresponds to $f$. By definition the lowest level mapped to by $\phi$ is $f_{min}$, so $\phi$ maps the elements of $T_b$ to elements of $T_p^{k+1}$ between levels $f_{min}$ and $f(l)$ inclusive. In fact, it maps $T_b$ into a copy of $T_p^{f(l)-f_{min}+1}$ defined as the elements in the down-set of $\phi(l)$ that are in levels $f_{min}$ to $f(l)$ of $T^{k+1}$, inclusive. Call this

copy $T_f$. We can construct an embedding $\psi$ of $T_b+$ into $T_p^{k+1}$ as follows. Choose some integer $i$ between 1 and $f_{min} - 1$; this is the number of levels by which we will "push down" the embedding $\phi$ so as to "fit in" the element $m$. (So, if $f_{min} = 1$ this construction does not yield an embedding of $T_b+$, which agrees with $\mu(f)\beta(f) = 0$ for $f_{min} = 1$.) Define $\psi(l)$ to be $\phi(l)$ and define $\psi(m)$ to be any element in level $f(l) - i$ that is below $\psi(l)$. Once this choice is made $\psi$ is then determined. Consider the copy of $T_p^{f(l)-i}$ that is the down-set of $\psi(m)$. By the choice of $i$, this has at least as many levels as $T_f$, so just considering the top $f(l) - f_{min} + 1$ levels, we have a copy of $T_f$. Then, for all $x \in T_b+$ with $x \neq l, m$, define $\psi(x)$ to be the element in this copy of $T_f$ that corresponds to the element $\phi(x)$ in the original $T_f$. Since for each $i$ we have a choice of $p^i$ elements for $\psi(m)$, the total number of distinct embeddings this construction yields for a particular $\phi$ that corresponds to $f$ is

$$\sum_{i=1}^{f_{min}-1} p^i = \frac{p^{f_{min}} - p}{p - 1} = \beta(f)$$

Since there are $\mu(f)$ distinct embeddings that correspond to $f$, this construction yields $\sum_{f \in \mathcal{F}''} \mu(f)\beta(f)$ distinct embeddings of $T_b+$ into $T_p^{k+1}$ that map $l$ to an element in level $k$ or $k + 1$.

Since each embedding of $T_b+$ into $T_p^{k+1}$ that maps $l$ to level $k$ or $k + 1$ can be converted to an embedding of $T_b$ into $T_p^{k+1}$ that maps $l$ to level $k$ or $k+1$ by reversing the above construction, we have that the total number of embeddings of $T_b+$ into $T_p^{k+1}$ that map $l$ to an element in level $k$ or $k + 1$ is exactly $\sum_{f \in \mathcal{F}''} \mu(f)\beta(f)$. Therefore, $\sum_{f \in \mathcal{F}''} \mu(f)\beta(f) = pd_k + d_{k+1}$ and $\sum_{f \in \mathcal{F}''} \mu(f)\alpha(f)\beta(f) = d_{k+1}$. So, applying Theorem 8.1 gives $c_{k+1}(pd_k + d_{k+1}) \leq (pc_k + c_{k+1})d_{k+1}$ which is equivalent to the inequality $d_k/c_k \leq d_{k+1}/c_{k+1}$.

So, we have two increasing sequences $(b_k/a_k)$ and $(d_k/c_k)$ for $k = 1, \ldots, n$. We need to apply Theorem 8.1 once more to a very simple lattice, namely the $n$-element chain, $[n]$. A chain is obviously a distributive lattice, and moreover *any* function $\mu$ is log-supermodular, since $\{k, k'\} = \{k \wedge k', k \vee k'\}$ for all $k, k' \in [n]$. Define $\mu(k) = a_k c_k$, define $\alpha(k) = b_k/a_k$, and define $\beta(k) = d_k/c_k$. Then $\alpha$ and $\beta$ are

increasing on $[n]$, and applying Theorem 8.1 gives

$$\sum_{k=1}^{n} b_k c_k \sum_{k=1}^{n} a_k d_k \leq \sum_{k=1}^{n} a_k c_k \sum_{k=1}^{n} b_k d_k. \tag{8.6}$$

Note that this inequality is the weighted version of the elementary inequality known as Chebyshev's Sum Inequality (see, for example, [16, Theorem 43]).

But $\sum_{k=1}^{n} a_k c_k$ is the total number of embeddings of $T_1$ into $T_p^n$, as we split $T_1$ into $T_t$ and $T_b$. Similarly, $\sum_{k=1}^{n} a_k d_k$ is the total number of embeddings of $T_2$ into $T_p^n$, as we split $T_2$ into $T_t$ and $T_b+$. Since $b_k$ only counts those embeddings counted by $a_k$ that also map the root of $T_t$ to $1_n$, we have that $\sum_{k=1}^{n} b_k c_k$ is the number of embeddings of $T_1$ into $T_p^n$ that map the root of $T_1$ to $1_n$, and $\sum_{k=1}^{n} b_k d_k$ is the number of embeddings of $T_2$ into $T_p^n$ that map the root of $T_2$ to $1_n$.

Therefore equation (8.6) becomes

$$A_{T_1}^{(p)}(n) C_{T_2}^{(p)}(n) \leq C_{T_1}^{(p)}(n) A_{T_2}^{(p)}(n)$$

as required.    □

Note that the proof is similar in its approach to the original proof by Kubicki, Lehel and Morayne; however in the set-up where we can apply the FKG-inequality we can view this result as one of many possible correlation inequalities on the lattice $\mathcal{F}(n; T)$, for $T$ some binary tree. Informally, in the proof of Theorem 8.4 we first show that the events "the root of $T_t$ is mapped to a high level of $T_p^n$" and "the element $l$ is mapped to a high level of $T_p^n$" are positively correlated on the lattice $\mathcal{F}(n; T_t)$. We then show that in the lattice $\mathcal{F}(k; T_b)$ having "$l$ mapped to a high level of $T_p^k$" means "the number of ways to embed an extra element" increases. We combine these correlations to show that if the root of $T_1$ is embedded "higher up" in $T_p^n$, then there are more embeddings of an extra element into $T_p^n$.

We can use the lattice $\mathcal{F}(n; T)$ and the function $\mu$ and other pairs of increasing functions on $\mathcal{F}$, to find other correlation inequalities. For example, we have the following result, which informally says that for any binary tree $T$ and any two

elements $x, y$ in $T$, the events "$x$ is mapped to a high level of $T_p^n$" and "$y$ is mapped to a high level of $T_p^n$" are positively correlated.

**Theorem 8.5.** *For any binary tree $T$, and any elements $x, y \in T$, and for any $k$ and $l$ with $1 \leq k, l < n$, we have*

$$\frac{E(k+1, l)}{E(k, l)} \leq \frac{E(k+1, l+1)}{E(k, l+1)},$$

*where $E(i, j)$ is the number of embeddings of $T$ into $T_p^n$ that map $x$ into level $i$, and $y$ into level $j$.*

**Proof.** Consider the sublattice $\mathcal{F}'$ of $\mathcal{F}(n; T)$ defined by $\mathcal{F}' = \{f \in \mathcal{F}(n; T) : f(x) = k, k+1 \text{ and } f(y) = l, l+1\}$. We take $\mu$ to be our log-supermodular function as described above, so that $\sum_{f \in \mathcal{F}'} \mu(f)$ is exactly

$$E(k, l) + E(k+1, l) + E(k, l+1) + E(k+1, l+1).$$

Define $\alpha(f) = I\{f(x) = k+1\}$ as the indicator of the event $f(x) = k+1$, and define $\beta(f) = I\{f(y) = l+1\}$ as the indicator of the event $f(y) = l+1$. Both $\alpha$ and $\beta$ are increasing on $\mathcal{F}'$ and so we can apply Theorem 8.1. This gives the inequality

$$[E(k+1, l) + E(k+1, l+1)] [E(k, l+1) + E(k+1, l+1)]$$

$$\leq [E(k, l) + E(k+1, l) + E(k, l+1) + E(k+1, l+1)] E(k+1, l+1)$$

which is equivalent to the required inequality.            □

This statement is not true if $T$ is allowed to be arbitrary, as illustrated by the following example. Let $T$ be a tree with 4 elements, the root $x$ and its three lowers covers $x_1, x_2, x_3$. Suppose we are embedding $T$ into $T^4$, the complete binary tree on 4 levels. We can calculate the different number of embeddings that map the elements $x_1$ and $x_2$ into particular levels. There are 12 embeddings that map $x_1$ to level 3 and $x_2$ to level 2, there are 32 embeddings that map $x_1$ to level 3 and $x_2$ to level 1, there are 76 embeddings that map $x_1$ to level 2 and $x_2$ to level 2 and there are 184 embeddings that map $x_1$ to level 2 and $x_2$ to level 1. So, if we consider a

uniform probability distribution over all embeddings of $T$ into $T^n$, we have that the conditional probability that an embedding maps $x_2$ into level 2, given that it maps $x_2$ into either level 1 or 2 and maps $x_1$ into level 3, is $12/44 = 3/11$. However, the conditional probability that an embedding maps $x_2$ into level 2, given that it maps $x_2$ into either level 1 or 2 and maps $x_1$ into level 2, is $76/260 = 19/65$ which is greater than $3/11$. In other words, it is more likely for $x_2$ to be in the higher of the two levels 1 and 2, if $x_1$ is in the lower of the two levels 2 and 3. This is still true for embeddings of $T$ into $T_p^4$ for $p > 2$. This means that we are unable to use this approach even for embeddings of $p$-ary trees into the complete $p$-ary tree.

## 8.2 Order-preserving maps of arbitrary trees into the complete $p$-ary tree

We can consider the case of $T$ being binary as special. For arbitrary $T$ we cannot define a log-supermodular function $\mu$ on $\mathcal{F}(n; T)$ so that $\sum_{f \in \mathcal{F}(n;T)} \mu(f)$ is the number of embeddings of $T$ into $T_p^n$. However, we can look at other types of mapping from $T$ into $T_p^n$, for example order-preserving maps. Recall that an order-preserving map preserves comparability of elements, but may introduce extra relations between elements. We look at both *strict* and *weak* order-preserving maps, the difference essentially being that a strict order-preserving map must map comparable elements to distinct elements, but a weak order-preserving map need not. We give formal definitions later.

### 8.2.1 Strict order-preserving maps

For strict order-preserving maps, the situation is very much simplified; as we have seen in the proof of Theorem 8.4 the function $\mu_1$, which counts the number of strict order-preserving maps, is log-supermodular with equality on $\mathcal{F}$. Moreover, if we

allow $T$ to be arbitrary, the function $\mu_1$ still counts the number of strict order-preserving maps. This is essentially because a strict order-preserving map only needs to preserve edges and not incomparability between elements. Therefore we can generalise the correlation inequalities for embeddings of binary trees to correlation inequalities for strict-order preserving maps of arbitrary trees.

Recall that a strict order-preserving map is a map $\phi$ from $T$ to $T_p^n$ such that $x > y$ in $T$ implies $\phi(x) > \phi(y)$ in $T_p^n$. Define $\bar{A}_T^{(p)}(n)$ to be the number of strict order-preserving maps of $T$ into $T_p^n$ that map the root of $T$ to $1_n$, and define $\bar{C}_T^{(p)}(n)$ to be the total number of strict order-preserving maps of $T$ into $T_p^n$. We have the following result, corresponding to the inequality of Theorem 8.4.

**Theorem 8.6.** *If $T_1$ and $T_2$ are trees such that $T_2$ contains a subposet isomorphic to $T_1$, then*

$$\frac{\bar{A}_{T_1}^{(p)}(n)}{\bar{C}_{T_1}^{(p)}(n)} \leq \frac{\bar{A}_{T_2}^{(p)}(n)}{\bar{C}_{T_2}^{(p)}(n)}$$

**Proof.** We follow the proof method of Theorem 8.4, making the necessary changes for strict order-preserving maps of arbitrary trees.

Firstly, note that we can define a distributive lattice of level functions $\mathcal{F}(n;T)$ when $T$ is an arbitrary tree. We take $\mu_1$ defined, as before, as

$$\mu_1(f) = p^{n-f(1_T)} \prod_{x>y,\text{ an edge in } T} p^{f(x)-f(y)},$$

which is a log-supermodular function. This satisfies log-supermodularity with equality (as noted in the proof of Theorem 8.4). Also, for any tree $T$, the sum

$$\sum_{f \in \mathcal{F}(n;T)} \mu_1(f)$$

is the number of strict order-preserving maps of $T$ into $T_p^n$, as explained earlier.

As before, we can assume that $T_1$ is isomorphic to the subposet $T_2 \setminus \{m\}$ of $T_2$, where $m$ is some element of $T_2$. Let $l$ be the upper cover of $m$ in $T_2$. We split $T_1$ into $T_t = T_1 \setminus D(l)$ and $T_b = D[l]$ as a subposet of $T_1$, and split $T_2$ into $T_t$ and $T_b+ = T_b \cup \{m\}$.

Set $\mathcal{F}' = \{f \in \mathcal{F}(n; T_t) : f(l) = k \text{ or } f(l) = k+1\}$ for $1 \leq k < n$ and let $\alpha(f) = I\{f(1_{T_t}) = n\}$ and $\beta(f) = I\{f(l) = k+1\}$, which are both increasing on $\mathcal{F}'$.

For $k = 1, \ldots, n$, define $\bar{a}_k$ to be the number of strict order-preserving maps of $T_t$ into $T_p^n$ that map $l$ to an element of level $k$, and define $\bar{b}_k$ to be the number of strict order-preserving maps of $T_t$ into $T_p^n$ that map $l$ to an element of level $k$ and map the root of $T_t$ to the root $1_n$. Here, we have

$$\sum_{f \in \mathcal{F}'} \mu_1(f)\alpha(f) = \bar{b}_k + \bar{b}_{k+1}, \qquad\qquad \sum_{f \in \mathcal{F}'} \mu_1(f) = \bar{a}_k + \bar{a}_{k+1},$$

$$\sum_{f \in \mathcal{F}'} \mu_1(f)\beta(f) = \bar{a}_{k+1}, \qquad\qquad \sum_{f \in \mathcal{F}'} \mu_1(f)\alpha(f)\beta(f) = \bar{b}_{k+1},$$

and applying Theorem 8.1 we get

$$\frac{\bar{b}_k}{\bar{a}_k} \leq \frac{\bar{b}_{k+1}}{\bar{a}_{k+1}},$$

in a similar way as in the proof of Theorem 8.4.

Now we look at trees $T_b$ and $T_b+$ and define $\bar{c}_k$ to be the number of strict order-preserving maps of $T_b$ into $T_p^k$ that map $l$ to $1_k$, and define $\bar{d}_k$ to be the number of strict order-preserving maps of $T_b+$ into $T_p^k$ that map $l$ to $1_k$. Whereas in the proof of Theorem 8.4 we had two cases to consider (from the two cases in Lemma 8.2), here, since the trees $T_1, T_2$ are not necessarily binary, we cannot be so specific. However, we just need that $m$ is the lower cover of $l$ in $T_b+$, where $l$ is the root of $T_b+$.

Let $\mathcal{F}'' = \{f \in \mathcal{F}(k+1; T_b) : f(l) = k \text{ or } f(l) = k+1\}$ for $1 \leq k < n$ and let $\alpha(f) = I\{f(l) = k+1\}$. Recall that $D(m)$ is the set of elements below $m$ in $T_b+$. We can consider $D(m)$ as a subposet of either $T_b$ or $T_b+$. Let $\beta(f) = (p^{\bar{f}_{min}} - p)/(p-1)$, where $\bar{f}_{min} = \min_{x \in D(m) \cup \{l\}} f(x)$. We have that $\alpha$ and $\beta$ are increasing on $\mathcal{F}''$. As before, the sum $\sum_{f \in \mathcal{F}''} \mu_1(f)$ equals $p\bar{c}_k + \bar{c}_{k+1}$ and the sum $\sum_{f \in \mathcal{F}''} \mu_1(f)\alpha(f)$ equals $\bar{c}_{k+1}$.

We now show that $\sum_{f \in \mathcal{F}''} \mu_1(f)\beta(f) = p\bar{d}_k + \bar{d}_{k+1}$ and $\sum_{f \in \mathcal{F}''} \mu_1(f)\alpha(f)\beta(f) = \bar{d}_{k+1}$. Note that $D[m]$, the subtree of $T_b+$ of elements below or equal to $m$ in $T_b+$

is isomorphic to the subtree $D(m) \cup \{l\}$ of $T_b$. In a similar way as in the proof of Theorem 8.4 we construct strict order-preserving maps from $T_b+$ to $T_p^{k+1}$ using strict order-preserving maps from $T_b$ to $T_p^{k+1}$. Fix $f$ in $\mathcal{F}''$ and let $\phi$ be a strict order-preserving map from $T_b$ to $T_p^{k+1}$ that corresponds to $f$. By definition of $\bar{f}_{min}$, the map $\phi$ maps the elements of $D(m) \cup \{l\}$ to elements of $T_p^{k+1}$ between levels $\bar{f}_{min}$ and $f(l)$ inclusive. So, it maps $D(m) \cup \{l\}$ into a copy of $T_p^{f(l)-\bar{f}_{min}+1}$ defined as the elements in the down set of $\phi(l)$ that are in levels $\bar{f}_{min}$ to $f(l)$ of $T_p^{k+1}$, inclusive. Call this copy $T_f$. We construct $\psi$ a strict order-preserving map from $T_b+$ to $T_p^{k+1}$ as follows. For all $x \in T_b \setminus D(m)$ set $\psi(x) = \phi(x)$. Choose some integer $i$ between 1 and $\bar{f}_{min} - 1$. Define $\psi(m)$ to be any element in level $f(l) - i$ that is below $\psi(l)$. Since we are constructing an order-preserving map, it does not matter if we choose an element that is comparable, or even equal to $\psi(x)$ for some $x \in T_b \setminus (D(m) \cup \{l\})$. So, we have a choice of $p^i$ elements. Once the choice is made $\psi$ is then determined. Consider the down-set of $\psi(m)$, which is a copy of $T_p^{f(l)-i}$. By the choice of $i$, this has a least as many levels as $T_f$, so considering just the top $f(l) - \bar{f}_{min} + 1$ levels we have a copy of $T_f$. Then, for all $x \in D(m)$, define $\psi(x)$ to be the element in this copy of $T_f$ that corresponds to the element $\phi(x)$ in the original $T_f$. Each choice of $i$ and choice of element $\psi(m)$ gives a distinct strict order-preserving map from $T_b+$ to $T_p^{k+1}$, so this construction yields

$$\sum_{i=1}^{\bar{f}_{min}-1} p^i = \frac{p^{\bar{f}_{min}} - p}{p-1} = \beta(f)$$

distinct strict order-preserving maps for a particular $\phi$ that corresponds to $f$. There are $\mu_1(f)$ distinct strict order-preserving maps that correspond to $f$, each yielding $\beta(f)$ distinct strict order-preserving maps, so we can construct a total of

$$\sum_{f \in \mathcal{F}''} \mu_1(f)\beta(f)$$

distinct strict order-preserving maps from $T_b+$.

Since each strict order-preserving map from $T_b+$ to $T_p^{k+1}$ can be converted to a strict order-preserving map from $T_b$ to $T_p^{k+1}$ by reversing the above construction,

and the level that $l$ is mapped to is unchanged in the construction, we have that $\sum_{f \in \mathcal{F}''} \mu_1(f)\beta(f)$ is the total number of strict order-preserving maps from $T_b+$ to $T_p^{k+1}$ that map $l$ to an element in level $k$ or $k+1$. Therefore, $\sum_{f \in \mathcal{F}''} \mu_1(f)\beta(f) = p\bar{d}_k + \bar{d}_{k+1}$ and $\sum_{f \in \mathcal{F}''} \mu_1(f)\alpha(f)\beta(f) = \bar{d}_{k+1}$, and applying Theorem 8.1 gives the required inequality $\bar{d}_k/\bar{c}_k \leq \bar{d}_{k+1}/\bar{c}_{k+1}$.

Finally, as in the proof of Theorem 8.4, we have increasing sequences $(\bar{b}_k/\bar{a}_k)$ and $(\bar{d}_k/\bar{c}_k)$ and a final application of Theorem 8.1 gives

$$\sum_{k=1}^{n} \bar{b}_k \bar{c}_k \sum_{k=1}^{n} \bar{a}_k \bar{d}_k \leq \sum_{k=1}^{n} \bar{a}_k \bar{c}_k \sum_{k=1}^{n} \bar{b}_k \bar{d}_k.$$

which, by inspection of each sum, is identical to the inequality

$$\bar{A}_{T_1}^{(p)}(n)\bar{C}_{T_2}^{(p)}(n) \leq \bar{C}_{T_1}^{(p)}(n)\bar{A}_{T_2}^{(p)}(n)$$

as required. $\square$

As with embeddings of binary trees, by applying the FKG-inequality to different increasing functions, versions of this proof can be used to establish other correlation inequalities for strict order-preserving maps of arbitrary trees into the complete $p$-ary tree.

## 8.2.2 Weak order-preserving maps

We have an analogous result for weak order-preserving maps from $T$ to $T_p^n$. A *weak order-preserving map* is a map $\phi$ from $T$ to $T_p^n$ such that $x > y$ in $T$ implies $\phi(x) \geq \phi(y)$ in $T_p^n$. Note that a function which maps all of $T$ to a single element of $T_p^n$ is a weak order-preserving map.

Define $\bar{\bar{A}}_T^{(p)}(n)$ to be the number of weak order-preserving maps of $T$ into $T_p^n$ that map $1_T$ to $1_n$, and define $\bar{\bar{C}}_T^{(p)}(n)$ to be the total number of weak order-preserving maps of $T$ into $T_p^n$.

We have the corresponding inequality as follows.

**Theorem 8.7.** *If $T_1$ and $T_2$ are trees such that $T_2$ contains a subposet isomorphic to $T_1$, then*

$$\frac{\bar{A}_{T_1}^{(p)}(n)}{\bar{\bar{C}}_{T_1}^{(p)}(n)} \leq \frac{\bar{A}_{T_2}^{(p)}(n)}{\bar{\bar{C}}_{T_2}^{(p)}(n)}.$$

**Proof.** The proof is naturally very similar to that for strict order-preserving maps. We follow that proof through, highlighting the differences for weak order-preserving maps.

For any tree $T$, write $\bar{\mathcal{F}}(n;T)$ for the lattice of weak order preserving maps from a tree $T$ to $[n]$. So, $f \in \bar{\mathcal{F}}(n;T)$ is a function from $T$ to $[n]$ such that $x > y$ in $T$ implies $f(x) \geq f(y)$ in $[n]$. As for $\mathcal{F}(n;T)$, the lattice of strict order-preserving maps from $T$ to $[n]$, the ordering on $\bar{\mathcal{F}}(n;T)$ is $f \geq g$ if and only if $f(x) \geq g(x)$ for all $x \in T$. Again, the join, $f \vee g$, is the pointwise maximum of $f$ and $g$, and the meet, $f \wedge g$, is the pointwise minimum of $f$ and $g$, and we have that $\bar{\mathcal{F}}(n;T)$ is a distributive lattice.

We call a function in $\bar{\mathcal{F}}(n;T)$ a *weak level function*. Every weak order-preserving map $\phi$ from $T$ to $T_p^n$ corresponds to a weak level function $f$ by setting $f(x)$ equal to the level of $\phi(x)$ in $T_p^n$. Moreover, if $\mu_1$ is defined on $\bar{\mathcal{F}}(n;T)$ as

$$\mu_1(f) = p^{n-f(1_T)} \prod_{x > y, \text{ an edge in } T} p^{f(x)-f(y)}$$

then $\sum_{f \in \bar{\mathcal{F}}(n;T)} \mu_1(f)$ is equal to $\bar{\bar{C}}_T^{(p)}(n)$ the number of weak order-preserving maps from $T$ to $T_p^n$.

As before, the function $\mu_1$ is log-supermodular (with equality) on $\bar{\mathcal{F}}(n;T)$.

Assume $T_1$ is isomorphic to $T_2 \setminus \{m\}$, for some $m \in T_2$. As in the earlier proofs, we split $T_1$ into $T_t$ and $T_b$, and split $T_2$ into $T_t$ and $T_b+$. Let $\bar{\mathcal{F}}' = \{f \in \bar{\mathcal{F}}(n;T_t) : f(l) = k \text{ or } f(l) = k+1\}$ and let $\alpha(f) = I\{f(1_{T_t}) = n\}$ and $\beta(f) = I\{f(l) = k+1\}$ which are both increasing on $\bar{\mathcal{F}}'$.

We define $\bar{a}_k$ to be the number of weak order-preserving maps of $T_t$ into $T_p^n$ that map $l$ to an element of level $k$, and define $\bar{b}_k$ to be the number of weak order-

preserving maps of $T_t$ into $T_p^n$ that map $l$ to an element of level $k$ and map the root of $T_t$ to the root $1_n$. Then, as in the proof of Theorem 8.6 we apply Theorem 8.1 to $\bar{\mathcal{F}}', \mu_1, \alpha, \beta$ to get

$$\frac{\bar{\bar{b}}_k}{\bar{\bar{a}}_k} \leq \frac{\bar{\bar{b}}_{k+1}}{\bar{\bar{a}}_{k+1}}.$$

Now, define $\bar{c}_k$ to be the number of weak order-preserving maps of $T_b$ into $T_p^k$ that map $l$ to $1_k$, and define $\bar{d}_k$ to be the number of weak order-preserving maps of $T_b+$ into $T_p^k$ that map $l$ to $1_k$. Let $\bar{\mathcal{F}}'' = \{f \in \bar{\mathcal{F}}(k+1; T_b) : f(l) = k \text{ or } f(l) = k+1\}$ and let $\alpha(f) = I\{f(l) = k+1$ and $\beta(f) = (p^{\bar{f}_{min}} - 1)/(p-1)$ where $\bar{f}_{min} = \min_{x \in D(m) \cup \{l\}} f(x)$.

Given a weak order-preserving map from $T_b$ to $T_p^{k+1}$ we use the same construction as described in the proof of Theorem 8.6 to construct weak order-preserving maps from $T_b+$ to $T_p^{k+1}$. However, note that a weak order-preserving map from $T_b+$ is allowed to map the elements $l$ and $m$ to the same element in $T_p^{k+1}$. In order to also construct these maps we allow the choice for $i$ to include 0, so that the level we pick for the element $m$ can be the same as the level for $l$. Therefore, for a particular weak order-preserving map $\phi$ from $T_b$ to $T_p^{k+1}$ corresponding to some $f$, our construction yields

$$\sum_{i=0}^{\bar{f}_{min}} p^i = \frac{p^{\bar{f}_{min}} - 1}{p - 1} = \beta(f)$$

distinct weak order-preserving maps from $T_b+$ to $T_p^{k+1}$. This means that

$$\sum_{f \in \bar{\mathcal{F}}''} \mu_1(f)\beta(f) = p\bar{\bar{d}}_k + \bar{\bar{d}}_{k+1} \quad \text{and} \quad \sum_{f \in \bar{\mathcal{F}}''} \mu_1(f)\alpha(f)\beta(f) = \bar{\bar{d}}_{k+1},$$

as in the proof of Theorem 8.6.

So, we apply Theorem 8.1, giving

$$\frac{\bar{\bar{d}}_k}{\bar{\bar{c}}_k} \leq \frac{\bar{\bar{d}}_{k+1}}{\bar{\bar{c}}_{k+1}}.$$

Finally, as before, we can apply Theorem 8.1 a final time, to the sequences

$(\bar{\bar{b}}_k/\bar{\bar{a}}_k)$ and $(\bar{\bar{d}}_k/\bar{\bar{c}}_k)$ to get

$$\sum_{k=1}^{n} \bar{\bar{b}}_k \bar{\bar{c}}_k \sum_{k=1}^{n} \bar{\bar{a}}_k \bar{\bar{d}}_k \le \sum_{k=1}^{n} \bar{\bar{a}}_k \bar{\bar{c}}_k \sum_{k=1}^{n} \bar{\bar{b}}_k \bar{\bar{d}}_k.$$

which by inspection of each sum, is identical to the inequality

$$\bar{\bar{A}}_{T_1}^{(p)}(n) \bar{C}_{T_2}^{(p)}(n) \le \bar{C}_{T_1}^{(p)}(n) \bar{\bar{A}}_{T_2}^{(p)}(n)$$

as required.                                                                     □

## 8.3   Related open problems

We finish this chapter by stating some open problems.

We have shown that Conjecture 4.4 does not hold for arbitrary trees, and we have the result of Theorem 4.1 for binary trees. Does the inequality hold for other trees? Our counterexamples in Section 6.4 show that we cannot allow arbitrary ternary trees. However, all our counterexamples have the property that $l(T_1) = l(T_2)$ and $d(T_1) = d(T_2)$; recall that Corollary 6.5 implies that this is a necessary condition for the pair of trees to be an asymptotic counterexample. Could it be that if either (i) $l(T_1) < l(T_2)$, or (ii) $l(T_1) = l(T_2)$ and $d(T_1) < d(T_2)$, then the trees $T_1, T_2$ satisfy the inequality?

**Question 8.8.** *Is it the case that, for any $n$ and any trees $T_1, T_2$ with $T_1$ a subposet of $T_2$ and either*

*(i) $l(T_1) < l(T_2)$, or*

*(ii) $l(T_1) = l(T_2)$ and $d(T_1) < d(T_2)$,*

*we have*

$$\frac{A_{T_1}^{(2)}(n)}{C_{T_1}^{(2)}(n)} \le \frac{A_{T_2}^{(2)}(n)}{C_{T_2}^{(2)}(n)}?$$

Alternatively, we could restrict to the case where $T_1$ is obtained from $T_2$ by only removing leaves. Again, this would exclude all of the counterexamples presented earlier. From experience we believe that disallowing these pairs of trees, where the extra element of $T_2$ is not a leaf, is enough to imply the inequality. Unfortunately, when we remove the restriction on $T_1$ and $T_2$ being binary, we are no longer able to apply the FKG-inequality and we are back to looking for a brute-force counting argument. We believe we have such an argument for ternary trees $T_1, T_2$, but this method will not generalise to arbitrary trees.

**Conjecture 8.9.** *For any $n$ and any trees $T_1, T_2$ such that an isomorphic copy of $T_1$ can be obtained by sequentially removing leaves from $T_2$, we have*

$$\frac{A_{T_1}^{(2)}(n)}{C_{T_1}^{(2)}(n)} \leq \frac{A_{T_2}^{(2)}(n)}{C_{T_2}^{(2)}(n)}.$$

# Chapter 9

# FKG-type inequalities for product lattices

In Chapter 8 we used the FKG-inequality to prove correlation inequalities for certain maps of trees into complete trees. We were able to find a distributive lattice $\mathcal{F}$ with a log-supermodular function $\mu$ and increasing functions $\alpha$ and $\beta$ on $\mathcal{F}$, so that the sums in the FKG-inequality (8.1) counted the specific mappings we were interested in. What if we have $\mathcal{F}, \mu, \alpha, \beta$ so that the sums are of interest, but we do not have increasing functions $\alpha, \beta$? In this chapter we show that it is sometimes possible to get a correlation inequality

$$\sum_{f \in \mathcal{F}} \mu(f)\alpha(f) \sum_{f \in \mathcal{F}} \mu(f)\beta(f) \leq \sum_{f \in \mathcal{F}} \mu(f) \sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\beta(f) \qquad (9.1)$$

like the FKG-inequality, even if one of the functions $\alpha$ or $\beta$ is not increasing.

To be precise, we consider the case when $\mathcal{F}$ is a product lattice $\mathcal{T} \times \mathcal{U}$, and $\beta$ is not increasing, but "tiered" on $\mathcal{T} \times \mathcal{U}$, meaning that for all $t_1 > t_2$ in $\mathcal{T}$, the minimum value of $\beta$ on $\{t_1\} \times \mathcal{U}$ is greater than or equal to the maximum value on $\{t_2\} \times \mathcal{U}$. This condition means that we can find closed intervals $I_t$ of $\mathbb{R}$ for each $t \in \mathcal{T}$, such that: (a) $\beta(t, u) \in I_t$ for all $u \in \mathcal{U}$, and (b) if $t_1 > t_2$ then the interval $I_{t_1}$ lies entirely to the right of the interval $I_{t_2}$ in $\mathbb{R}$ (allowing touching end-points).

The general idea is to "average out" the deviations within $I_t$ of $\beta(t, u)$ over $\{t\} \times \mathcal{U}$ for each $t \in \mathcal{T}$ to obtain an increasing function on $\mathcal{T} \times \mathcal{U}$, and to then apply the FKG-inequality. We define $\tilde{\beta}(t, u)$ to be the weighted average

$$\tilde{\beta}(t, u) = \frac{\sum_{u \in \mathcal{U}} \mu(t, u)\beta(t, u)}{\sum_{u \in \mathcal{U}} \mu(t, u)}$$

which, by construction, is constant on $\{t\} \times \mathcal{U}$ so that $\tilde{\beta}(t, u) = \tilde{\beta}(t)$. Now, by (a), $\tilde{\beta}(t)$ is some real in the interval $I_t$ and, by (b), we have $\tilde{\beta}(t_1) \geq \tilde{\beta}(t_2)$ for all $t_1 > t_2 \in \mathcal{T}$, so $\tilde{\beta}$ is increasing on $\mathcal{T} \times \mathcal{U}$.

So, if we also have the usual conditions that $\mu$ is log-supermodular on $\mathcal{T} \times \mathcal{U}$ and $\alpha$ is increasing on $\mathcal{T} \times \mathcal{U}$, then we can apply the FKG-inequality to the lattice $\mathcal{T} \times \mathcal{U}$ and functions $\mu$, $\alpha$ and $\tilde{\beta}$. By construction of the "average" function $\tilde{\beta}$, we have that the sums $\sum_{f \in \mathcal{F}} \mu(f)\tilde{\beta}(f)$ and $\sum_{f \in \mathcal{F}} \mu(f)\beta(f)$ are equal, and with the extra condition that $\alpha$ is constant on $\{t\} \times \mathcal{U}$, for all $t \in \mathcal{T}$, we have that the sums $\sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\tilde{\beta}(f)$ and $\sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\beta(f)$ are also equal and we get the correlation inequality (9.1) for $\alpha$ and $\beta$.

In fact, we can give a more general result, where $\mathcal{F}$ is a sublattice of a product lattice $\mathcal{T} \times \mathcal{U} \times \mathcal{V}$, and for each $v \in \mathcal{V}$ the function $\beta$ is tiered on $\mathcal{T} \times \mathcal{U} \times \{v\}$. In this situation, extra conditions are required to ensure that the method described above of "averaging out" $\beta$ still yields an increasing function $\tilde{\beta}$. It should also be pointed out that the condition on $\alpha$ is crucial; we have no other way of ensuring that $\sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\tilde{\beta}(f) = \sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\beta(f)$. Before stating the result we give the following definition.

**Definition 9.1.** For $\mathcal{F}$ a sublattice of $\mathcal{T} \times \mathcal{U} \times \mathcal{V}$ let $\mathcal{F}|_{t,v} = \{u \in \mathcal{U} : (t, u, v) \in \mathcal{F}\}$, for each $t \in \mathcal{T}, v \in \mathcal{V}$.

**Lemma 9.2.** *Suppose $\mathcal{F}$ is a sublattice of some product lattice $\mathcal{T} \times \mathcal{U} \times \mathcal{V}$ and $\mu, \alpha, \beta$ are non-negative functions defined on $\mathcal{F}$, with $\mu$ log-supermodular on $\mathcal{F}$ and $\alpha$ increasing on $\mathcal{F}$. If we have the further conditions,*

(1) *for all $t_1 > t_2 \in \mathcal{T}$, $v \in \mathcal{V}$ with $\mathcal{F}|_{t_1,v} \neq \emptyset$ and $\mathcal{F}|_{t_2,v} \neq \emptyset$,*

$$\min_{u \in \mathcal{F}|_{t_1,v}} \beta(t_1, u, v) \geq \max_{u \in \mathcal{F}|_{t_2,v}} \beta(t_2, u, v),$$

(2) $\mathcal{F}|_{t,v_1} = \mathcal{F}|_{t,v_2}$ *for all $t \in \mathcal{T}$, $v_1, v_2 \in \mathcal{V}$ with $\mathcal{F}|_{t,v_1} \neq \emptyset$ and $\mathcal{F}|_{t,v_2} \neq \emptyset$,*

(3) *if $(t, u, v_1) \geq (t, u, v_2) \in \mathcal{F}$ then $\beta(t, u, v_1) \geq \beta(t, u, v_2)$,*

(4) $\mu(t, u, v) = \mu_1(t, u)\mu_2(t, v)$ *for some $\mu_1, \mu_2$,*

(5) $\alpha$ *is constant on $\{t\} \times \mathcal{F}|_{t,v} \times \{v\}$ for all $t \in \mathcal{T}, v \in \mathcal{V}$,*

*then the FKG-inequality holds. That is,*

$$\sum_{f \in \mathcal{F}} \mu(f)\alpha(f) \sum_{f \in \mathcal{F}} \mu(f)\beta(f) \leq \sum_{f \in \mathcal{F}} \mu(f) \sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\beta(f).$$

Let us informally discuss these conditions. As described earlier, the idea is to "average out" the deviations of $\beta$ over the $\mathcal{U}$-coordinate. Note that conditions (1) and (3) imply that $\beta$ is increasing in the $\mathcal{T}$- and $\mathcal{V}$-coordinates, so we just need to perform the average in a way that preserves this monotonicity. Condition (1) means that for each $t \in \mathcal{T}$, $v \in \mathcal{V}$ with $\mathcal{F}|_{t,v}$ non-empty we can find a closed interval $I_{t,v}$ of $\mathbb{R}$ such that $\beta(t, u, v) \in I_{t,v}$ for all $u \in F|_{t,v}$ and the interval $I_{t_1,v}$ is entirely to the right of the interval $I_{t_2,v}$ in $\mathbb{R}$ if $t_1 > t_2$ in $\mathcal{T}$. (Note that we assume nothing on the ordering of intervals $I_{t_1,v_1}$ and $I_{t_2,v_2}$ for $v_1 \neq v_2$.) So, for each $t \in \mathcal{T}, v \in \mathcal{V}$ with $\mathcal{F}|_{t,v}$ non-empty, we can average $\beta$ over $\{t\} \times \mathcal{F}|_{t,v} \times \{v\}$ to obtain a new function $\tilde{\beta}$ defined as a weighted average

$$\tilde{\beta}(t, u, v) = \frac{\sum_{u \in \mathcal{F}|_{t,v}} \lambda(t, u, v)\beta(t, u, v)}{\sum_{u \in \mathcal{F}|_{t,v}} \lambda(t, u, v)}$$

and condition (1) ensures that this $\tilde{\beta}$ is increasing in the $\mathcal{T}$-coordinate. By construction, $\tilde{\beta}$ is constant on $\mathcal{F}|_{t,v}$ for all $t \in \mathcal{T}, v \in \mathcal{V}$, so we just need to ensure $\tilde{\beta}$ is increasing in the $\mathcal{V}$-coordinate. One way of achieving this is to take the same weights $\lambda$ for different $v \in \mathcal{V}$, that is, to assume $\lambda$ is just a function of $t$ and $u$, and

to assume conditions (2) and (3). Then the function

$$\tilde{\beta}(t, u, v) = \frac{\sum_{u \in \mathcal{F}|_{t,v}} \lambda(t, u)\beta(t, u, v)}{\sum_{u \in \mathcal{F}|_{t,v}} \lambda(t, u)}$$

will be increasing in the $\mathcal{V}$-coordinate. Finally, we assume condition (4) on $\mu$ so that we can find an appropriate weight $\lambda(t, u)$ that ensures that the sums $\sum_{f \in \mathcal{F}} \mu(f)\tilde{\beta}(f)$ and $\sum_{f \in \mathcal{F}} \mu(f)\beta(f)$ are equal. (We will see that this weight $\lambda(t, u)$ should be the factor $\mu_1(t, u)$.) We assume condition (5) which, with condition (4), ensures that the sums $\sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\tilde{\beta}(f)$ and $\sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\beta(f)$ are equal.

These conditions may seem rather arbitrary, and it is reasonable to ask whether we can find examples of lattices and functions satisfying them. We will show later in this chapter, that the lattice of level functions studied in the previous chapter, with some familiar functions, do satisfy the conditions of Lemma 9.2 and this will enable us to give an alternative proof of one of the cases of Theorem 8.4.

Before proving Lemma 9.2, we state some corollaries, which are special cases of the lemma and follow immediately by interpreting conditions (1)–(5) for the particular case. If $\mathcal{F}$ is the whole lattice $\mathcal{T} \times \mathcal{U} \times \mathcal{V}$, then we have the following, since $\mathcal{F}|_{t,v} = \mathcal{U}$ for all $t \in \mathcal{T}, v \in \mathcal{V}$.

**Corollary 9.3.** *Suppose $\mathcal{F}$ is a product lattice $\mathcal{F} = \mathcal{T} \times \mathcal{U} \times \mathcal{V}$ and $\mu, \alpha, \beta$ are non-negative functions defined on $\mathcal{F}$, with $\mu$ log-supermodular on $\mathcal{F}$ and $\alpha$ increasing on $\mathcal{F}$. If we have the further conditions,*

(1) *for all $t_1 > t_2 \in \mathcal{T}$, $v \in \mathcal{V}$,*

$$\min_{u \in \mathcal{U}} \beta(t_1, u, v) \geq \max_{u \in \mathcal{U}} \beta(t_2, u, v),$$

(2) *if $(t, u, v_1) \geq (t, u, v_2) \in \mathcal{F}$ then $\beta(t, u, v_1) \geq \beta(t, u, v_2)$,*

(3) *$\mu(t, u, v) = \mu_1(t, u)\mu_2(t, v)$ for some $\mu_1, \mu_2$,*

(4) *$\alpha$ is constant on $\{t\} \times \mathcal{U} \times \{v\}$ for all $t \in \mathcal{T}, v \in \mathcal{V}$,*

*then the FKG-inequality holds. That is,*

$$\sum_{f \in \mathcal{F}} \mu(f)\alpha(f) \sum_{f \in \mathcal{F}} \mu(f)\beta(f) \leq \sum_{f \in \mathcal{F}} \mu(f) \sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\beta(f). \qquad \square$$

Also as a corollary to Lemma 9.2 is the case where $\mathcal{V}$ is a single element, so that $\mathcal{T} \times \mathcal{U} \times \mathcal{V}$ can be thought of as the product $\mathcal{T} \times \mathcal{U}$. For $\mathcal{F}$ a sublattice of $\mathcal{T} \times \mathcal{U}$, write $\mathcal{F}|_t$ for the set $\{u \in \mathcal{U} : (t, u) \in \mathcal{F}\}$.

**Corollary 9.4.** *Suppose $\mathcal{F}$ is a sublattice of some product lattice $\mathcal{T} \times \mathcal{U}$ and $\mu, \alpha, \beta$ are non-negative functions defined on $\mathcal{F}$, with $\mu$ log-supermodular on $\mathcal{F}$ and $\alpha$ increasing on $\mathcal{F}$. If we have the further conditions,*

*(1) for all $t_1 > t_2 \in \mathcal{T}$ with $\mathcal{F}|_{t_1} \neq \emptyset$ and $\mathcal{F}|_{t_2} \neq \emptyset$,*

$$\min_{u \in \mathcal{F}|_{t_1}} \beta(t_1, u) \geq \max_{u \in \mathcal{F}|_{t_2}} \beta(t_2, u),$$

*(2) $\alpha$ is constant on $\{t\} \times \mathcal{F}|_t$ for all $t \in \mathcal{T}$,*

*then the FKG-inequality holds. That is,*

$$\sum_{f \in \mathcal{F}} \mu(f)\alpha(f) \sum_{f \in \mathcal{F}} \mu(f)\beta(f) \leq \sum_{f \in \mathcal{F}} \mu(f) \sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\beta(f). \qquad \square$$

Note that the case of $\mathcal{F}$ being the whole lattice $\mathcal{T} \times \mathcal{U}$ is the special case informally described at the beginning of the chapter. For completeness, we state the formal result here.

**Corollary 9.5.** *Suppose $\mathcal{F}$ is a product lattice $\mathcal{F} = \mathcal{T} \times \mathcal{U}$ and $\mu, \alpha, \beta$ are non-negative functions defined on $\mathcal{F}$, with $\mu$ log-supermodular on $\mathcal{F}$ and $\alpha$ increasing on $\mathcal{F}$. If we have the further conditions,*

*(1) for all $t_1 > t_2 \in \mathcal{T}$,*

$$\min_{u \in \mathcal{U}} \beta(t_1, u) \geq \max_{u \in \mathcal{U}} \beta(t_2, u),$$

*(2) $\alpha$ is constant on $\{t\} \times \mathcal{U}$ for all $t \in \mathcal{T}$,*

*then the FKG-inequality holds. That is,*

$$\sum_{f \in \mathcal{F}} \mu(f)\alpha(f) \sum_{f \in \mathcal{F}} \mu(f)\beta(f) \le \sum_{f \in \mathcal{F}} \mu(f) \sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\beta(f). \qquad \square$$

We now give the proof of Lemma 9.2.

**Proof.** Define

$$\tilde{\beta}(t,v) = \sum_{u \in \mathcal{F}|_{t,v}} \mu_1(t,u)\beta(t,u,v) \Big/ \sum_{u \in \mathcal{F}|_{t,v}} \mu_1(t,u) \qquad (9.2)$$

for all $t \in \mathcal{T}, v \in \mathcal{V}$ with $\mathcal{F}|_{t,v} \ne \emptyset$. We can think of $\tilde{\beta}$ as a function on $\mathcal{F}$, by defining $\tilde{\beta}(t,u,v) = \tilde{\beta}(t,v)$ for all $(t,u,v) \in \mathcal{F}$. We have that $\tilde{\beta}$ is increasing on $\mathcal{F}$, as follows. For $(t_1,u,v) > (t_2,u,v) \in \mathcal{F}$, so that $\mathcal{F}|_{t_1,v} \ne \emptyset$ and $\mathcal{F}|_{t_2,v} \ne \emptyset$, we have

$$\tilde{\beta}(t_1,v) = \sum_{u \in \mathcal{F}|_{t_1,v}} \mu_1(t_1,u)\beta(t_1,u,v) \Big/ \sum_{u \in \mathcal{F}|_{t_1,v}} \mu_1(t_1,u)$$

$$\ge \min_{u \in \mathcal{F}|_{t_1,v}} \beta(t_1,u,v) \ge \max_{u \in \mathcal{F}|_{t_2,v}} \beta(t_2,u,v) \qquad \text{by (1)}$$

$$\ge \sum_{u \in \mathcal{F}|_{t_2,v}} \mu_1(t_2,u)\beta(t_2,u,v) \Big/ \sum_{u \in \mathcal{F}|_{t_2,v}} \mu_1(t_2,u) = \tilde{\beta}(t_2,v).$$

For $(t,u,v_1) \ge (t,u,v_2) \in \mathcal{F}$, so that $\mathcal{F}|_{t,v_1} \ne \emptyset$ and $\mathcal{F}|_{t,v_2} \ne \emptyset$, we have $\mathcal{F}|_{t,v_1} = \mathcal{F}|_{t,v_2}$ by (2). So,

$$\tilde{\beta}(t,v_1) = \sum_{u \in \mathcal{F}|_{t,v_1}} \mu_1(t,u)\beta(t,u,v_1) \Big/ \sum_{u \in \mathcal{F}|_{t,v_1}} \mu_1(t,u)$$

$$= \sum_{u \in \mathcal{F}|_{t,v_2}} \mu_1(t,u)\beta(t,u,v_1) \Big/ \sum_{u \in \mathcal{F}|_{t,v_2}} \mu_1(t,u)$$

$$\ge \sum_{u \in \mathcal{F}|_{t,v_2}} \mu_1(t,u)\beta(t,u,v_2) \Big/ \sum_{u \in \mathcal{F}|_{t,v_2}} \mu_1(t,u) = \tilde{\beta}(t,v_2),$$

where the inequality follows from (3). Since $\tilde{\beta}$ is, by definition, independent of the $\mathcal{U}$-coordinate, and increasing in the $\mathcal{T}$- and $\mathcal{V}$-coordinates it is increasing on $\mathcal{F}$.

So, $\alpha$ and $\tilde{\beta}$ are increasing on $\mathcal{F}$ and $\mu$ is log-supermodular and we can apply the FKG-inequality, giving

$$\sum_{f \in \mathcal{F}} \mu(f)\alpha(f) \sum_{f \in \mathcal{F}} \mu(f)\tilde{\beta}(f) \le \sum_{f \in \mathcal{F}} \mu(f) \sum_{f \in \mathcal{F}} \mu(f)\alpha(f)\tilde{\beta}(f)$$

It just remains to show that the sums $\sum_{f \in \mathcal{F}} \mu(f) \tilde{\beta}(f)$ and $\sum_{f \in \mathcal{F}} \mu(f) \alpha(f) \tilde{\beta}(f)$ are equal to the sums $\sum_{f \in \mathcal{F}} \mu(f) \beta(f)$ and $\sum_{f \in \mathcal{F}} \mu(f) \alpha(f) \beta(f)$.

We have

$$
\sum_{(t,u,v) \in \mathcal{F}} \mu(t, u, v) \tilde{\beta}(t, v) = \sum_{\substack{(t,v) \in \mathcal{T} \times \mathcal{V} \\ \mathcal{F}|_{t,v} \neq \emptyset}} \sum_{u \in \mathcal{F}|_{t,v}} \mu_1(t, u) \mu_2(t, v) \tilde{\beta}(t, v) \qquad \text{by (4)}
$$

$$
= \sum_{\substack{(t,v) \in \mathcal{T} \times \mathcal{V} \\ \mathcal{F}|_{t,v} \neq \emptyset}} \mu_2(t, v) \left( \tilde{\beta}(t, v) \sum_{u \in \mathcal{F}|_{t,v}} \mu_1(t, u) \right)
$$

$$
= \sum_{\substack{(t,v) \in \mathcal{T} \times \mathcal{V} \\ \mathcal{F}|_{t,v} \neq \emptyset}} \mu_2(t, v) \sum_{u \in \mathcal{F}|_{t,v}} \mu_1(t, u) \beta(t, u, v) \qquad \text{by (9.2)}
$$

$$
= \sum_{(t,u,v) \in \mathcal{F}} \mu(t, u, v) \beta(t, u, v).
$$

By (5), we can view $\alpha$ as a function $\alpha(t, v)$ of just the $\mathcal{T}$- and $\mathcal{V}$-coordinates, so we have

$$
\sum_{(t,u,v) \in \mathcal{F}} \mu(t, u, v) \alpha(t, v) \tilde{\beta}(t, v) = \sum_{\substack{(t,v) \in \mathcal{T} \times \mathcal{V} \\ \mathcal{F}|_{t,v} \neq \emptyset}} \sum_{u \in \mathcal{F}|_{t,v}} \mu_1(t, u) \mu_2(t, v) \alpha(t, v) \tilde{\beta}(t, v)
$$

$$
= \sum_{\substack{(t,v) \in \mathcal{T} \times \mathcal{V} \\ \mathcal{F}|_{t,v} \neq \emptyset}} \mu_2(t, v) \alpha(t, v) \left( \tilde{\beta}(t, v) \sum_{u \in \mathcal{F}|_{t,v}} \mu_1(t, u) \right)
$$

$$
= \sum_{\substack{(t,v) \in \mathcal{T} \times \mathcal{V} \\ \mathcal{F}|_{t,v} \neq \emptyset}} \mu_2(t, v) \alpha(t, v) \sum_{u \in \mathcal{F}|_{t,v}} \mu_1(t, u) \beta(t, u, v)
$$

$$
= \sum_{(t,u,v) \in \mathcal{F}} \mu(t, u, v) \alpha(t, v) \beta(t, u, v)
$$

which completes the proof.                                                                $\square$

We finish by using the above results to give an alternative proof to Theorem 8.4 in the case where the trees $T_1$ and $T_2$ differ by one element $m$, a leaf of $T_2$, and the upper cover of $m$ is not a leaf in $T_1$. Recall in the earlier proof, that in this case we did not use the FKG-inequality to show that the sequence $d_k/c_k$ is increasing.

**Alternative proof of a case of Theorem 8.4.** Let $T_1$ and $T_2$ be binary trees

with $T_2 = T_1 \setminus \{m\}$ for some leaf $m$ of $T_2$. We will show that

$$\frac{A_{T_1}^{(p)}(n)}{C_{T_1}^{(p)}(n)} \le \frac{A_{T_2}^{(p)}(n)}{C_{T_2}^{(p)}(n)}$$

As in chapter 8 we work with a lattice of level functions. Let $l$ be the upper cover of $m$ and let $m_1$ be the lower cover of $l$ that is different from $m$. Recall that $D(m_1)$ is the set of all elements in $T_1$ that are below $m_1$. We work with the lattice of level functions $\mathcal{F}(n; T_1 \setminus D(m_1))$. Let $h$ be the height of $D[m_1] = D(m_1) \cup \{m_1\}$, and let $\mathcal{F}'$ be the sublattice $\mathcal{F}' = \{f \in \mathcal{F}(n; T_1 \setminus D(m_1)) : f(m_1) \ge h\}$. As before, we have a log-supermodular function $\mu$ on $\mathcal{F}'$ defined as

$$\mu(f) = p^{n-f(1_{T_1})} \prod_{\substack{x > y, \text{ an edge} \\ \text{in } T_1 \setminus D(m_1)}} p^{f(x)-f(y)} \prod_{\substack{y \in T_1 \setminus D(m_1), \\ y \text{ has 2 lower} \\ \text{covers, } z_1, z_2}} (1 - p^{\max\{f(z_1), f(z_2)\}-f(y)}).$$

Let $\alpha(f) = I\{f(1_{T_1}) = n\}$ and let $\beta(f) = \sum_{i=f(m)}^{f(l)-1}(p^i - 1)/(p - 1)$.

Note that, since $h$ is the height of $D[m_1]$, any embedding $\psi$ of $T_1$ into $T_p^n$ must map $m_1$ into level $h$ or higher. This means that the level function $g$ corresponding to $\psi$ has $g(m_1) \ge h$, so $g$ restricted to the set $T_1 \setminus D(m_1)$ is in $\mathcal{F}'$. That is, the restriction of any embedding of $T_1$ into $T_p^n$ to the set $T_1 \setminus D(m_1)$ yields an embedding of $T_1 \setminus D(m_1)$ into $T_p^n$ that corresponds to some level function in $\mathcal{F}'$.

Conversely, for each embedding $\phi$ of $T_1 \setminus D(m_1)$ into $T_p^n$ that corresponds to $f \in \mathcal{F}'$, we can construct $A = A_{D[m_1]}^{(p)}(f(m_1))$ embeddings $\psi_i$ for $i = 1, \ldots, A$ of $T_1$ into $T_p^n$, as follows. Write $\theta_i$, $i = 1, \ldots, A$ for the distinct embeddings of $D[m_1]$ into $T_p^{f(m_1)}$ that map $m_1$ to $1_{f(m_1)}$. Since $\phi(m_1)$ is an element in level $f(m_1)$ of $T_p^n$, the down-set of $\phi(m_1)$ is a copy of $T_p^{f(m_1)}$. So, for $x \in T_1 \setminus D(m_1)$, define $\psi_i(x) = \phi(x)$, and for $x \in D(m_1)$ define $\psi_i(x)$ to be the element in this copy of $T_p^{f(m_1)}$ that corresponds the the element $\theta_i(x)$. We have that $\psi_i$ is an embedding of $T_1$ into $T_p^n$, and the $\psi_i$ are all distinct. Since there are $\mu(f)$ embeddings of $T_1 \setminus D(m_1)$ into $T_p^n$ that correspond to $f$ we have a total of $\mu(f)A_{D[m_1]}^{(p)}(f(m_1))$ distinct embeddings of $T_1$ into $T_p^n$ for each $f \in \mathcal{F}$. Therefore, $C_{T_1}^{(p)}(n) = \sum_{f \in \mathcal{F}} \mu(f)A_{D[m_1]}^{(p)}(f(m_1))$.

Notice that, since $\{f, g\} = \{f \wedge g, f \vee g\}$ for all $f, g \in \mathcal{F}$, we have

$$A^{(p)}_{D[m_1]}(f(m_1))A^{(p)}_{D[m_1]}(g(m_1)) = A^{(p)}_{D[m_1]}((f \wedge g)(m_1))A^{(p)}_{D[m_1]}((f \vee g)(m_1))$$

so that the function $\mu'(f) = \mu(f)A^{(p)}_{D[m_1]}(f(m_1))$ is also log-supermodular on $\mathcal{F}'$, and we have

$$C^{(p)}_{T_1}(n) = \sum_{f \in \mathcal{F}} \mu'(f), \qquad A^{(p)}_{T_1}(n) = \sum_{f \in \mathcal{F}} \mu'(f)\alpha(f).$$

We now show

$$C^{(p)}_{T_2}(n) = \sum_{f \in \mathcal{F}} \mu'(f)\beta(f), \qquad A^{(p)}_{T_2}(n) = \sum_{f \in \mathcal{F}} \mu'(f)\alpha(f)\beta(f),$$

as follows.

As before, the restriction of any embedding of $T_2$ into $T^n_p$ to the set $T_1 \setminus D(m_1)$ yields an embedding of $T_1 \setminus D(m_1)$ that corresponds to some level function in $\mathcal{F}$. We show that for each $f \in \mathcal{F}'$, we can construct $\mu'(f)\beta(f)$ embeddings of $T_2$ into $T^n_p$. For each $f \in \mathcal{F}'$ we can construct $\mu'(f)$ embeddings of $T_1$ into $T^n_p$ using the construction described above. Let $\phi$ be such an embedding. We construct an embedding $\psi$ of $T_2$ into $T^n_p$ by setting $\psi(x) = \phi(x)$ for all $x \in T_1$, and choosing an element of $T^n_p$ for $\psi(m)$. We require $\psi(m)$ to be below $\psi(l)$ but incomparable with $\psi(m_1)$ and since $\psi(l) = \phi(l)$ is in level $f(l)$ of $T^n_p$ and $\psi(m_1) = \phi(m_1)$ is in level $f(m_1)$ we have a choice of $\sum^{f(l)-1}_{i=f(m_1)}(p^i - 1)/(p - 1) = \beta(f)$ elements for $\psi(m)$. Note that by the regularity of $T^n_p$ the number of places only depends on the level function and not the exact positions of $\phi(l)$ and $\phi(m_1)$. See Figure 9.1, for an example where $p = 2$, $f(l) = 5$ and $f(m_1) = 2$. Clearly each choice defines a different embedding of $T_2$ into $T^n_p$, so the total number of embeddings of $T_2$ into $T^n_p$ is $\sum_{f \in \mathcal{F}} \mu'(f)\beta(f)$.

Furthermore, if we have $f \in \mathcal{F}'$ with $f(1_{T_1 \setminus D(m_1)}) = n$, the construction yields an embedding that maps the root $1_{T_2} = 1_{T_1 \setminus D(m_1)}$ of $T_2$ to the root $1_n$ of $T^n_p$. Therefore, we have $A^{(p)}_{T_2}(n) = \sum_{f \in \mathcal{F}} \mu'(f)\alpha(f)\beta(f)$, as claimed.

It remains to show the inequality (9.1). We would like to apply the FKG-inequality, but the function $\beta$ is not increasing. However, we see that the dominant

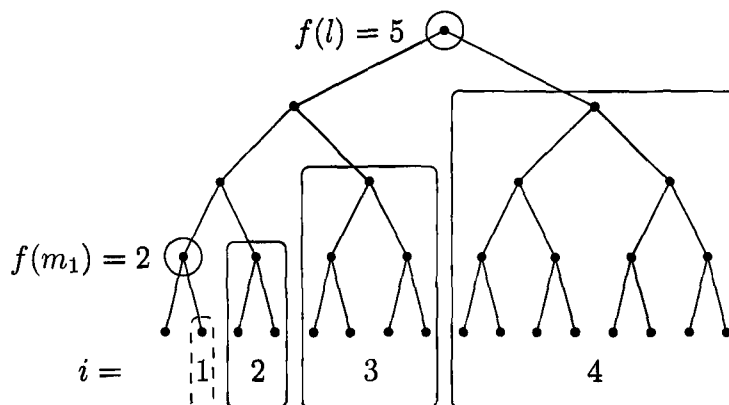$$f(l) = 5$$

$$f(m_1) = 2$$

$$i =$$ 1   2   3   4

Figure 9.1: The number of places to map $m$ is $\beta(f) = \sum_{i=f(m_1)}^{f(l)-1}(p^i - 1)/(p - 1)$

term in the sum $\sum_{i=f(m_1)}^{f(l)-1}(p^i - 1)/(p - 1)$ is the last term; moreover, it is larger than the sum of all the previous terms. This means that if we have $f, g \in \mathcal{F}'$ with $f(l) > g(l)$ then $\beta(f) \geq \beta(g)$ whatever the values of $f(m_1)$ and $g(m_1)$. This appears very similar to condition (1) of Lemma 9.2 and we now show that we can apply the lemma.

Suppose $l$ is not the root of $T_1$. Let $\mathcal{T} \times \mathcal{U} \times \mathcal{V}$ be the product lattice $[n] \times [n] \times \mathcal{F}(n; T_1 \setminus D[l])$. So an element $(t, u, v)$ is a triple whose first two coordinates are elements in $[n]$, and the third is a level function of $T_1 \setminus D[l]$. Let $k$ be the upper cover of $l$, and recall that $h$ is the height of $D[m_1]$. Consider the following sublattice $\{(t, u, v) \in \mathcal{T} \times \mathcal{U} \times \mathcal{V} : v(k) > t > u \geq h\}$ of $\mathcal{T} \times \mathcal{U} \times \mathcal{V}$. For each element $(t, u, v)$ in this sublattice we can define a function $f : T_1 \setminus D(m_1) \to \mathbb{R}$ as

$$f(x) = \begin{cases} v(x) & \text{for } x \in T_1 \setminus D[l], \\ t & \text{for } x = l, \\ u & \text{for } x = m_1. \end{cases}$$

and since $v(k) > t > u \geq h$ we have $f(k) > f(l) > f(m_1) \geq h$ which means that $f$ is in $\mathcal{F}'$. Conversely, for each level function $f \in \mathcal{F}'$ we can define a triple

$(t, u, v) \in \mathcal{T} \times \mathcal{U} \times \mathcal{V}$ by

$$t = f(l),$$

$$u = f(m_1),$$

$$v = f|_{T_1 \setminus D[l]},$$

and since $v(k) > t > u \geq h$, by definition, we have that $(t, u, v)$ is in the sublattice $\{(t, u, v) \in \mathcal{T} \times \mathcal{U} \times \mathcal{V} : v(k) > t > u \geq h\}$.

So, we can think of $\mathcal{F}'$ as a sublattice of product lattice $\mathcal{T} \times \mathcal{U} \times \mathcal{V}$ by considering $f$ as the triple $(f(l), f(m_1), f|_{T_1 \setminus D[l]})$. The functions $\mu'$, $\alpha$ and $\beta$ are non-negative functions of $\mathcal{F}'$, and $\mu'$ is log-supermodular on $\mathcal{F}'$, and $\alpha$ is increasing on $\mathcal{F}'$, so we need to check that conditions (1)–(5) hold in order to apply Lemma 9.2.

For $t \in \mathcal{T}$ and $v \in \mathcal{V}$, the set $\mathcal{F}'|_{t,v}$ is non-empty when $v(k) > t > h$ and in this case we have $\mathcal{F}'|_{t,v} = \{h, h+1, \ldots, t-1\}$. We have $\beta(t, u, v) = \sum_{i=u}^{t-1} (p^i - 1)/(p-1)$. Suppose $t_1 > t_2 \in [n]$, $v \in \mathcal{F}(n; T_1 \setminus D[l])$ with $v(k) > t_1 > h$ and $v(k) > t_2 > h$. Then, since $t_1 - 1 \geq t_2$ and $p \geq 2$, we have

$$\min_{u \in \{h, \ldots, t_1 - 1\}} \sum_{i=u}^{t_1-1} (p^i - 1)/(p-1) = \sum_{i=t_1-1}^{t_1-1} (p^i - 1)/(p-1)$$

$$= (p^{t_1-1} - 1)/(p-1)$$

$$\geq (p^{t_2} - 1)/(p-1)^2$$

$$= \sum_{i=0}^{t_2-1} p^i/(p-1) \geq \max_{u \in \{h, \ldots, t_2-1\}} \sum_{i=u}^{t_2-1} (p^i - 1)/(p-1)$$

which means that condition (1) holds. Suppose we have $t \in [n]$, $v_1, v_2 \in \mathcal{F}(n; T_1 \setminus D[l])$ with $v_1(k) > t > h$ and $v_2(k) > t > h$. Then $\mathcal{F}'|_{t,v_1} = \{h, \ldots, t-1\} = \mathcal{F}'|_{t,v_2}$, so condition (2) holds. Since $\beta(t, u, v)$ is not dependent on $v$ we have condition (3). Using the definition of $\mu(f)$ on $\mathcal{F}'$, we have

$$\mu(f) = p^{n-f(1_{T_1})} \prod_{\substack{x > y, \text{ an edge} \\ \text{in } T_1 \setminus D(m_1)}} p^{f(x)-f(y)} \prod_{\substack{y \in T_1 \setminus D(m_1), \\ y \text{ has 2 lower} \\ \text{covers, } z_1, z_2}} (1 - p^{\max\{f(z_1), f(z_2)\} - f(y)}).$$

Consider the contribution of the element $m_1$ in the above expression. There is a factor of $p^{f(l)-f(m_1)}$ which appears in the first product because of the edge $l > m_1$.

There is no contribution to the second product since $m_1$ is not the lower cover of an element that has two lower covers (that is, $l$ does not have two lower covers). So, we can write the above as

$$\mu(f) = p^{f(l)-f(m_1)}p^{n-f(1_{T_1})} \prod_{\substack{x>y, \text{ an edge} \\ \text{in } T_1\backslash D[m_1]}} p^{f(x)-f(y)} \prod_{\substack{y\in T_1\backslash D[m_1], \\ y \text{ has 2 lower} \\ \text{covers, } z_1, z_2}} (1 - p^{\max\{f(z_1),f(z_2)\}-f(y)})$$

$$= p^{-f(m_1)}\left[p^{f(l)}p^{n-f(1_{T_1})} \prod_{\substack{x>y, \text{ an edge} \\ \text{in } T_1\backslash D[m_1]}} p^{f(x)-f(y)} \prod_{\substack{y\in T_1\backslash D[m_1], \\ y \text{ has 2 lower} \\ \text{covers, } z_1, z_2}} (1 - p^{\max\{f(z_1),f(z_2)\}-f(y)})\right].$$

Note the change in the subscript in the products to $T_1 \setminus D[m_1]$. So, writing $t = f(l)$, $u = f(m_1)$, $v = f|_{T_1\backslash D[l]}$, we have $\mu(t,u,v) = p^{-u}\mu_2(t,v)$ since the term in square brackets depends only on $t$ and $v$. Since $\mu'(t,u,v) = \mu(t,u,v)A^{(p)}_{D[m_1]}(u)$, we have $\mu'(t,u,v) = \mu_1(t,u)\mu_2(t,v)$, where $\mu_1(t,u) = p^{-u}A^{(p)}_{D[m_1]}(u)$ is a function just of $u$. Therefore condition (4) holds. Finally, we note that $\alpha(t,u,v) = I\{v(1_{T_1}) = n\}$ is not dependent on $u$, and so condition (5) holds, and applying Lemma 9.2 gives the result.

Recall that we assumed that $l$ was not the root of $T_1$. In the case where it is, then things simplify greatly, and the product lattice $\mathcal{T} \times \mathcal{U} \times \mathcal{V}$ reduces to the product $\mathcal{T} \times \mathcal{U} = [n] \times [n]$. The tree $T_1 \setminus D(m_1)$ is simply the 2-element chain $l > m_1$. We can think of the lattice of level functions $\mathcal{F}'$ as the sublattice $\{(t,u) \in \mathcal{T} \times \mathcal{U} : t > u \geq h\}$ of $\mathcal{T} \times \mathcal{U} = [n] \times [n]$ by considering a function $f \in \mathcal{F}'$ as the pair $(f(l), f(m_1))$. We want to apply Corollary 9.4 so we need to check conditions (1) and (2) hold. Since $\mathcal{F}'|_t$ is equal to $\{h, h+1, \ldots, t-1\}$ if $t > h$ and empty otherwise, and $\beta(t,u) = \sum_{i=u}^{t-1}(p^i - 1)/(p - 1)$ we have that if $t_1 > t_2$ then

$$\min_{u\in\{h,\ldots,t_1-1\}} \sum_{i=u}^{t_1-1}(p^i - 1)/(p - 1) \geq \max_{u\in\{h,\ldots,t_2-1\}} \sum_{i=u}^{t_2-1}(p^i - 1)/(p - 1)$$

exactly as before, so condition (1) holds. Also, the function $\alpha(t,u) = I\{t = n\}$ is independent of $u$ so condition (2) holds, and we can apply Corollary 9.4 which gives the result.                                                                □

# Bibliography

[1] M. H. Albert and A. M. Frieze, Random graph orders, Order **6** (1989), no. 1, 19–30.

[2] N. Alon, J. H. Spencer and P. Erdős, *The Probabilistic Method*, John Wiley & Sons, 1992.

[3] I. Anderson, *Combinatorics of Finite Sets*, Oxford University Press, 1987.

[4] S. Arya, M. J. Golin and K. Mehlhorn, On the expected depth of random circuits, Combin. Probab. Comput. **8** (1999), no. 3, 209–227.

[5] K. Azuma, Weighted sums of certain dependent random variables, Tôhoku Math. J. (2) **19** (1967), 357–367.

[6] B. Bollobás and G. Brightwell, Box-spaces and random partial orders, Trans. Amer. Math. Soc. **324** (1991), no. 1, 59–72.

[7] B. Bollobás and G. Brightwell, The width of random graph orders, Math. Sci. **20** (1995), no. 2, 69–90.

[8] B. Bollobás and G. Brightwell, The dimension of random graph orders, The mathematics of Paul Erdős, II, 51–69, Algorithms Combin., 14, Springer, Berlin, 1997.

[9] B. Bollobás and G. Brightwell, The structure of random graph orders, SIAM J. Discrete Math. **10** (1997), no. 2, 318–335.

[10] G. Brightwell, Models of random partial orders, Surveys in combinatorics, 1993 (Keele), 53–83, London Math. Soc. Lecture Note Ser., 187, Cambridge Univ. Press, Cambridge, 1993.

[11] G. Brightwell, personal communication.

[12] F. Dowker and S. Surya, Observables in extended percolation models of causal set cosmology, arXiv:gr-qc/0504069.

[13] T. S. Ferguson, Who solved the secretary problem?, Statist. Sci. 4 (1989), no. 3, 282–296.

[14] P. C. Fishburn, *Interval Orders And Interval Graphs*, John Wiley & Sons, 1985.

[15] J. P. Gilbert and F. Mosteller, Recognizing the maximum of a sequence, J. Amer. Statist. Assoc. **61** (1966), 35–73.

[16] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press, 1934, 1952.

[17] G. Kubicki, J. Lehel and M. Morayne, A ratio inequality for binary trees and the best secretary, Combin. Probab. Comput. **11** (2002), 149–161.

[18] G. Kubicki, J. Lehel and M. Morayne, An asymptotic ratio in the complete binary tree, Order **20** (2003), 91–97.

[19] G. Kubicki, J. Lehel and M. Morayne, Counting chains and antichains in the complete binary tree, Unpublished.

[20] H. Mahmoud and T. Tsukiji, A limit law for outputs in random recursive circuits, Algorithmica **31** (2001), no. 3, 403–412.

[21] H. Mahmoud and T. Tsukiji, On the internal structure of random recursive circuits, J. Comput. Appl. Math. **142** (2002), no. 1, 155–171.

[22] M. Morayne, Partial-order analogue of the secretary problem: The binary tree case. Discrete Math. **184** (1998), 165–181.

[23] B. Pittel and R. Tungol, A phase transition phenomenon in a random directed acyclic graph, Random Structures Algorithms **18** (2001), no. 2, 164–184.

[24] D. P. Rideout and R. D. Sorkin, Classical sequential growth dynamics for causal sets, Phys. Rev. D (3) **61** (2000), no. 2, 024002, 16pp.

[25] D. P. Rideout and R. D. Sorkin, Evidence for a continuum limit in causal set dynamics, Phys. Rev. D (3) **63** (2001), no. 10, 104011, 15pp.

[26] E. R. Scheinerman, A note on graphs and sphere orders, Journal of Graph Theory, Vol. 17, No. 3, 283–289 (1993).

[27] K. Simon, D. Crippa and F. Collenberg, On the distribution of the transitive closure in a random acyclic digraph, Algorithms—ESA '93 (Bad Honnef), 345–356, Lecture Notes in Comput. Sci., 726, Springer, Berlin, 1993.

[28] J. Spencer, Minimal scrambling sets of simple orders, Acta Math. Hungar. **22** (1971/72), 349–353.

[29] W. T. Trotter, *Combinatorics and Partially Ordered Sets: Dimension Theory*, The Johns Hopkins University Press, 1992.

[30] T. Tsukiji and F. Xhafa, On the depth of randomly generated circuits, Algorithms—ESA '96 (Barcelona), 208–220, Lecture Notes in Comput. Sci., 1136, Springer, Berlin, 1996.

[31] M. Varadarajan and D. P. Rideout, A general solution for classical sequential growth dynamics of causal sets, arXiv:gr-qc/0504066.

[32] N. C. Wormald, Differential equations of random processes and random graphs, Ann. Appl. Probab. **5** (1995), 1217–1235.

[33] N. C. Wormald, Analysis of greedy algorithms on graphs with bounded degrees. EuroComb '01 (Barcelona). Discrete Math. **273** (2003), no. 1–3, 235–260.