Continuum limits for classical sequential growth models

Graham Brightwell
Department of Mathematics
London School of Economics
Houghton Street
London WC2A 2AE
U.K.

Nicholas Georgiou
Department of Mathematics
University of Bristol
University Walk
Bristol BS8 1TW
U.K.

January 26, 2009

Abstract

A random graph order, also known as a transitive percolation process, is defined by taking a random graph on the vertex set $\{0, \dots, n-1\}$, and putting i below j if there is a path $i = i_1 \cdots i_k = j$ in the graph with $i_1 < \cdots < i_k$.

In [15], Rideout and Sorkin provide computational evidence that suitably normalised sequences of random graph orders have a "continuum limit". We confirm that this is the case, and show that the continuum limit is always a semiorder.

Transitive percolation processes are a special case of a more general class called classical sequential growth models. We give a number of results describing the large-scale structure of a general classical sequential growth model. We show that for any sufficiently large n, and any classical sequential growth model, there is a semiorder S on $\{0, \dots, n-1\}$ such that the random partial order on $\{0, \dots, n-1\}$ generated according to the model differs from S on an arbitrarily small proportion of pairs. We also show that, if any sequence of classical sequential growth models has a continuum limit, then this limit is (essentially) a semiorder. We give some examples of continuum limits that can occur.

Classical sequential growth models were introduced as the only models satisfying certain properties making them suitable as discrete models for space-time. Our results indicate that this class of models does not contain any that are good approximations to Minkowski space in any dimension ≥ 2.

1 Introduction

In [15], Rideout and Sorkin provide evidence for a "continuum limit of transitive percolation". A transitive percolation process, 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 $n = 1, 2, \dots$, 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. In the mathematics literature, $P_{n,p}$ is called a random graph order. These were introduced by Albert and Frieze [1] and have been studied further by Alon, Bollobás, Brightwell and

Janson [2], Bollobás and Brightwell [5, 6, 7], Kim and Pittel [11], Pittel and Tungol [13], and Simon, Crippa and Collenberg [16].

In this paper, 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 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 [4].

Definition 1.1. 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 $x_1, \dots, x_k$.)

Definition 1.2. For P a partially ordered measure space with probability measure μ, and k a natural number, define a *random sample* of k elements from P to be a sequence $x_1, \dots, x_k$ of elements of P, obtained by selecting k elements $x_1, \dots, x_k$ independently from P according to μ, and conditioning on the event that $x_1, \dots, x_k$ are distinct. A random sample can be thought of as a (random) finite partial order on the fixed ground-set $\{x_1, \dots, x_k\}$, inheriting the partial order from P.

For Q a finite partial order with ground-set labelled as $\{x_1, \dots, x_k\}$, and P a partially ordered measure space with measure μ, let $\lambda(Q; P)$ be the probability that the partial order inherited from P on a random sample $x_1, \dots, x_k$ 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 μ 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 $x_1, \dots, x_{|Q|}$ 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 1.3. A *continuum limit* of a sequence $(\mathcal{P}_n)_{n=1}^{\infty}$ of probability spaces, whose elements are finite posets, is an atomless partially ordered measure space P_∞ such that, for all finite labelled partial orders Q,

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

This notion of continuum limit is analogous to the notion of "graph limit" introduced by Lovász and Szegedy [12], and the subject of much subsequent interest.

In [15], 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 the sequence $\big(\mathbb{E}\lambda(Q; P_n)\big)_{n=1}^{\infty}$ 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 4, 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 1.4. *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) $\lim_{n\to\infty} (p^{-1} \log p^{-1}/n) = 0$,

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

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

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 S_c consists of the set $[0, 1]$, with the uniform measure on Borel sets, and the order $\prec$ given by $x \prec y$ if and only if $y - x > c$, which is a 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. 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 1 as an induced suborder. See Fishburn [8] for a proof of this and much more information about semiorders.

Transitive percolation processes form a one-parameter family of models from a larger family, called classical sequential growth models. These models, introduced by Rideout and Sorkin in [14], produce random partial orders sequentially. A particular *classical sequential growth model* $\mathcal{P}(\mathbf{t})$ is specified by a sequence $\mathbf{t} = (t_0, t_1, \dots)$ of non-negative constants, with either t_0 or t_1 non-zero. 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 $P_n = P_n(\mathbf{t})$. Often, we will be interested in a sequence of classical sequential growth models $\mathcal{P}_n$—more properly $\mathcal{P}(\mathbf{t}^{(n)})$—and an associated random sequence $P_n(\mathbf{t}^{(n)})$ of partial orders.

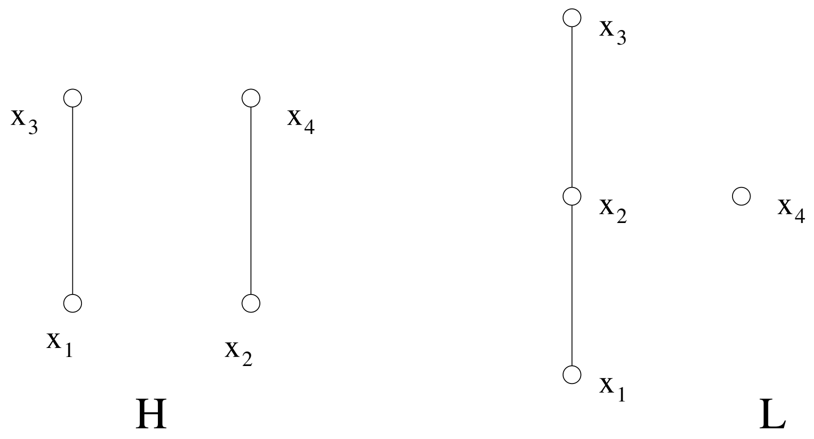

Figure 1: Forbidden induced suborders

Classical sequential growth models are of particular interest as they are the only ones satisfying some natural-looking conditions for discrete models of space-time—for further details, see [14].

A *transitive percolation process* with parameter p is a classical sequential growth model $\mathcal{P}(\mathbf{t})$, given by the sequence $\mathbf{t}$ with $t_i = t^i$ for all i, where $t = p/(1-p)$. Other examples of classical sequential growth models, where almost all the t_i are zero, include the "dust universe" ($t_i = 0$ for all $i \geq 1$), "forest universe" ($t_i = 0$ for all $i \geq 2$) and random binary growth model ($t_i = 0$ for all $i \geq 3$). The "dust" and "forest" universes are fairly trivial models; they produce an infinite antichain almost surely, and a infinite forest of infinite trees almost surely, respectively. The "dust" and "forest" universes are described in [14] and the random binary growth model has been studied by Georgiou [9].

It is natural to ask whether continuum limits exist for sequences of classical sequential growth models other than a sequence of transitive percolation processes, and in particular whether one can obtain continuum limits that are radically different from semiorders. In Section 5, we show that this is not possible.

Given two partial orders P, Q on the same finite ground set X, we define $\Delta(P, Q)$ to be the number of pairs of elements on which P and Q differ, i.e.,

$$\Delta(P, Q) = \#\{(x, y) \in X^{(2)} : P \text{ induces a different partial order from } Q \text{ on } \{x, y\}\}.$$

Theorem 1.5. *For any $\varepsilon > 0$ there exists an n_0 such that for any classical sequential growth model $\mathcal{P}(\mathbf{t})$ and for all $n \geq n_0$, there exists a semiorder S on $\{0, \ldots, n-1\}$ such that*

$$\mathbb{E}\Delta(P_n(\mathbf{t}), S) < \varepsilon n^2,$$

where $P_n(\mathbf{t})$ is the random order on $\{0, \ldots, n-1\}$ generated according to $\mathcal{P}(\mathbf{t})$.

This means that partial orders generated from classical sequential growth models resemble semiorders.

Note the independence of n_0 on $\mathbf{t}$ in the above result; this makes Theorem 1.5 applicable to a sequence of classical sequential growth models. Indeed, it follows from a rephrasing of the theorem in terms of random sampling that any sequence of classical sequential growth models $(\mathcal{P}_n)_{n=1}^{\infty}$ has $\mathbb{E}\lambda(H; P_n) \to 0$ and $\mathbb{E}\lambda(L; P_n) \to 0$ as $n \to \infty$ (where, as before, P_n is a random partial order produced by the model $\mathcal{P}_n$) and this motivates the following definition.

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, i.e., if $\lambda(H;P) = \lambda(L;P) = 0$. Following the above exposition we are able to classify the continuum limit of a sequence of classical sequential growth models, when it exists.

Theorem 1.6. *If a sequence of classical sequential growth models* $(\mathcal{P}_n)_{n=1}^{\infty}$ *has a continuum limit, then this limit is an almost-semiorder.*

This still leaves the question of whether continuum limits exist for sequences other than a sequence of transitive percolation processes. In Section 6, we show that they do and, in accordance with Theorem 1.6, their structure is a more general version of the continuum limit S_c described in Section 4.

Definition 1.7. For a Borel-measurable function $r : [0,1] \to [0,\infty]$, let T_r 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 μ_L is the Lebesgue measure on $[0,1]$, and $\prec$ is defined by $x \prec y$ if and only if $\int_x^y r(t)dt > 1$.

In particular, if r is constant, then $T_r = S_{\min\{1/r,1\}}$. Note that the T_r are all semiorders.

Consider a classical sequential growth model $\mathcal{P}(\mathbf{t})$ with ground set $\{0,\dots,n-1\}$. A key quantity is the sequence $\frac{\mathbb{E}|D_y|}{y}$ $(y = 1,\dots,n-1)$: for given y, this is the probability that each element x earlier than y is selected for D_y, and is thus the analogue of the constant p in transitive percolation. It turns out that the behaviour of this sequence is enough to determine the large-scale structure of $P_n(\mathbf{t})$, for large n. If we are interested in the structure of the partial order restricted to $[\varepsilon n, n-1]$, then Theorem 1.4 suggests that we should compare $\frac{y}{\mathbb{E}|D_y|}\log\left(y/\mathbb{E}|D_y|\right)$ with n. That is, for $\varepsilon n \le y \le n$, we should consider

$$r_n(y/n) = \frac{\mathbb{E}|D_y|}{(y/n)\log y}.$$

Extending this function to all real values $x \in [\varepsilon, 1]$ we set

$$r_n(x) = \frac{n\mathbb{E}|D_{\lceil xn\rceil}|}{\lceil xn\rceil \log\lceil xn\rceil} \sim \frac{\mathbb{E}|D_{\lceil xn\rceil}|}{x\log n}.$$

The next theorem can be interpreted as saying that $P_n(\mathbf{t})$ is close to T_{r_n}, in the sense of Theorem 1.5.

Theorem 1.8. *Suppose* $(\mathcal{P}_n)_{n=1}^{\infty}$ *is a sequence of classical sequential growth models with associated functions* r_n *as defined above. Suppose that* $r(x) = \lim_{n\to\infty} r_n(x) \in [0,\infty]$ *exists almost everywhere on* $[0,1]$. *Then* $(\mathcal{P}_n)_{n=1}^{\infty}$ *has a continuum limit, namely* T_r.

We shall give an example of a sequence of classical sequential growth models that satisfies the conditions of Theorem 1.8 and has a non-trivial continuum limit.

To a large extent, we describe the large-scale structure of an arbitrary classical sequential growth model, enabling us to answer questions about them arising in physics. Specifically, it has been asked [14, 15] whether classical sequential growth models can be constructed to resemble a "sprinkling" from Minkowski space M^d, for any dimension $d \ge 2$, i.e., a partial order obtained from M^d by taking points according to a Poisson process with fixed density λ. 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.

2 Preliminaries

We begin with a few preliminary observations about the probabilities $\lambda(Q;P)$.

Lemma 2.1. *For Q,P finite labelled partial orders with $|Q| = j$, and for k with $j \leq k \leq |P|$,*

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

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 $x_1 \ldots, x_k$ of k elements from P. The probability that the labelled subposet on $x_1,\ldots,x_j$ is equal to Q is the sum of $\lambda(Q',P)$ over all labelled partial orders Q' that, when restricted to $\{x_1,\ldots,x_j\}$, 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. □

Corollary 2.2. *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 2.1. □

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 simple result, applying to cases where the partial order is either very dense or very sparse.

Proposition 2.3.

(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 isomorphic to C_k. Define Q' as a relabelled copy of Q with the elements x_1, x_2 incomparable, which is possible since $Q \ncong C_k$. Note that $\lambda(Q';P_n) = \lambda(Q;P_n)$. Since A_2 is a subposet of Q', we can apply Corollary 2.2 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 isomorphic 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 $n \to \infty$. □

3 Random graph orders

We recall the definition of a random graph order.

Definition 3.1. 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: indeed, in the description above, we can think of the element n as being placed above a subset D with probability $p^{|D|}(1-p)^{n-|D|}$, which is proportional to $t^{|D|}$, where $t = p/(1-p)$, as in the earlier definition. In future, we will use the term random graph order, rather than transitive percolation process, but the terms are essentially interchangeable.

We include some results of Pittel and Tungol, from [13], which we will need in order to prove the existence of a continuum limit. Versions of this result can be found in earlier work of Simon, Crippa and Collenberg [16], and Bollobás and Brightwell [6]. We change the notation slightly, for ease of use in this paper. The following results apply to a random graph order $P_{N,\pi}$, and we will apply them with particular values for N and π. 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 3.2 (Pittel and Tungol, [13, 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 3.3 (Pittel and Tungol, [13, 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 3.4 (Pittel and Tungol, [13, 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

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

4 The continuum limits of $P_{n,p}$

We show that, for suitable functions $p(n)$, the continuum limit of the sequence $P_{n,p(n)}$ of random graph orders is the partially ordered measure space defined below, the order structure of which is a semiorder.

Definition 4.1. 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 μ_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 $x \not\prec y$ 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 4.2. *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, Proposition 2.3 implies that it is enough to show that $\mathbb{E}\lambda(A_2; P_n) \to 0$ as $n \to \infty$. Fix ε 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 3.2 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 3.2 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, Proposition 2.3 implies that it is enough to show that $\mathbb{E}\lambda(C_2; P_n) \to 0$ as $n \to \infty$. Fix ε 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 3.2 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 3.2 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 ε 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 $\mathbb{P}(\prec_n$ 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

$$\mathbb{P}(\prec_n \text{ induces different partial order from } \prec \text{ on } \{x,y\}) \to 0 \tag{1}$$

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 $q = \mathbb{P}(\prec_n$ induces different partial order from $\prec$ 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 (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 $c-\varepsilon \leq |i-j|/n \leq c+\varepsilon$ 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 and 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 ε-*bad* element if $|U[x] \cap [x + (c+\varepsilon)n]| < \varepsilon n/2$, and an ε-*good* element otherwise. We will show that the number of ε-bad elements is small, and the number of bad pairs (x,y) with x an ε-good element and $y - x > (c+\varepsilon)n$ is also small.

We can calculate the expected number of ε-bad elements as follows. Let $\pi = (1+\varepsilon/2c)\log{(c+\varepsilon)n}/(c+\varepsilon)n$. Since $p > \pi$, the expected number of ε-good elements in $P_{n,p}$ is greater than the expected number of ε-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 3.3 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

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

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 $\mathbb{P}(x \text{ is } \varepsilon\text{-bad in } P_{n,\pi}) \leq \mathbb{P}(\gamma^*_{(c+\varepsilon)n}(0) \leq M)$. Therefore, the probability that x is an ε-bad element in $P_{n,p}$ is $O(1/\log n) + o(1)$. So, the expected number of ε-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 ε-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 ε-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 ε-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 3.4 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 ε is arbitrary we have the result. □

To complete the proof of Theorem 1.4, 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 $0 \le c \le 1$. So, we have

Lemma 4.3. *For $0 \le c_1 \ne c_2 \le 1$, $\lambda(A_2; S_{c_1}) \ne \lambda(A_2; S_{c_2})$.* □

The following lemma is an obvious extension to Theorem 4.2 and is stated without proof.

Lemma 4.4. *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 4.2, then the subsequence has a continuum limit as described in Theorem 4.2.* □

Theorem 4.5. *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 4.2.*

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 4.4 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 $c_2 \ge 1$. In either case, by Lemma 4.3 we have $\lim_{n\to\infty} \mathbb{E}\lambda(A_2; P_{a_n,p}) \ne \lim_{n\to\infty} \mathbb{E}\lambda(A_2; P_{b_n,p})$. 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 1.4.

5 Continuum limits are almost-semiorders

In Section 4 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 $0 \le c \le 1$ 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.

The results we give apply to all sequences $\mathbf{t}$, 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 $\mathbf{t}$ that are delicately balanced. We do not wish to spend time here 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 4.2. 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 4.2, 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 4.2, and therefore we would expect the continuum limit of the growth model to be sparser than that of the random graph order, and hence be an antichain. See Section 6 for more details and an extension of Theorem 4.2.

Our task is to show that, for any continuum limit P_∞ of a sequence of classical sequential growth models, $\lambda(H; P_\infty) = \lambda(L; P_\infty) = 0$. Informally, we need to show that, for any classical sequential growth model $\mathcal{P}(\mathbf{t})$, the number of copies of H and L as subposets of $P_n(\mathbf{t})$ is small, for large enough n. To do this we study the large scale structure of $P_n(\mathbf{t})$. We prove Theorem 1.5, by showing that, according to some metric assigned to $P_n(\mathbf{t})$, elements that are more than distance 1 apart are likely to be comparable, and elements that are less than distance 1 apart are likely to be incomparable. This is a generalisation of the result for random graph orders, where the metric can be taken to be constant.

We begin with some lemmas describing some properties of classical sequential growth models. For now, we consider a fixed classical sequential growth model $\mathcal{P}(\mathbf{t})$, with terms t_i. 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 5.1. *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 $\mathbb{E}(|D_x|) \leq \mathbb{E}(|D_{x+1}|)$.

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

$$\mathbb{E}(|D_{x+1}|) \geq \frac{\sum_{j=0}^{x} j\binom{x+1}{j}t_j}{\sum_{j=0}^{x} \binom{x+1}{j}t_j}. \tag{2}$$

Now, note that $\binom{x+1}{j} = \frac{x+1}{x+1-j}\binom{x}{j}$ so the inequality (2) becomes

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

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., [10, Theorem 43]), since both j and $1/(x+1-j)$ are increasing on $\{0,1,\dots,x\}$. □

Our aim in the lemmas that follow is essentially to establish $\mathbb{E}|D_y|$, for y running from εn to n, as the key parameter determining how the model $\mathcal{P}(\mathbf{t})$ develops. Note that, in a transitive percolation process with parameter p, $\mathbb{E}|D_y|/y$ is constant throughout the process, and the size of D_y is concentrated around its mean. Of course, these properties do not hold for our general classical sequential growth model, so we will establish weaker results instead, that suffice for our purposes.

One possibility to bear in mind is that $|D_y|$ is, with very high probability, close to one value k, but the expectation of $|D_y|$ is much higher than k, because of the contribution of one term $m\binom{y}{m}t_m$, where m is much larger than k. So, at 'level' y, $\mathbb{E}|D_y|$ does not capture the behaviour of the process well. However, in this scenario, if we look instead at some value z only a little higher than y, we will discover that $\mathbb{E}|D_z|$ is already close to m, and indeed that it is very unlikely that $|D_z|$ is much smaller than m. As we are looking at the partial order on very coarse scales, the "transition phase" between y and z is immaterial.

The following lemma is the first step along this path. We show that, if the contribution to the expectation $\mathbb{E}|D_y|$ from terms greater than m is significant, then at some point z, not much bigger than y, the expectation $\mathbb{E}|D_z|$ is a constant fraction of m.

Lemma 5.2. *For $m, y \in \mathbb{N}$, $\gamma > 0$ and $0 < \eta < 1/2$, if*

$$\sum_{k=m}^{y}\binom{y}{k}kt_k \geq \gamma\sum_{k=0}^{y}\binom{y}{k}t_k,$$

then

$$\mathbb{E}|D_z| \geq m(1-2\eta)$$

for $z \in \mathbb{N}$ satisfying

$$y \leq z\left(1 - \frac{2\log(m/\gamma\eta)}{m\eta}\right).$$

Proof. Assume that $y \leq z\left(1 - 2\log(m/\gamma\eta)/m\eta\right)$. Working with $\left(\mathbb{E}|D_z|\right)^{-1}$, we have

$$\begin{aligned}\left(\mathbb{E}|D_z|\right)^{-1} = \frac{\sum_{k=0}^{z}\binom{z}{k}t_k}{\sum_{k=0}^{z}\binom{z}{k}kt_k} &\leq \frac{\sum_{k=0}^{m(1-\eta)}\binom{z}{k}t_k + \sum_{k=m(1-\eta)}^{z}\binom{z}{k}t_k}{\sum_{k=m(1-\eta)}^{z}\binom{z}{k}kt_k} \\ &\leq \frac{\sum_{k=0}^{m(1-\eta)}\binom{z}{k}t_k}{\sum_{k=m(1-\eta)}^{z}\binom{z}{k}kt_k} + \frac{1}{m(1-\eta)}.\end{aligned}$$

Note that $\binom{z}{l}/\binom{y}{l}$ is increasing in l, so that

$$\frac{\sum_{k=0}^{m(1-\eta)}\binom{z}{k}t_k}{\sum_{k=0}^{m(1-\eta)}\binom{y}{k}t_k} \leq \frac{\binom{z}{m(1-\eta)}}{\binom{y}{m(1-\eta)}} \qquad \text{and} \qquad \frac{\sum_{k=m}^{y}\binom{z}{k}kt_k}{\sum_{k=m}^{y}\binom{y}{k}kt_k} \geq \frac{\binom{z}{m}}{\binom{y}{m}}.$$

But

$$
\begin{aligned}
\frac{\binom{z}{m(1-\eta)}}{\binom{y}{m(1-\eta)}}\frac{\binom{y}{m}}{\binom{z}{m}} &= \frac{(y-m(1-\eta))!(z-m)!}{(z-m(1-\eta))!(y-m)!} \\
&= \frac{(y-m+1)\cdots(z-m)}{(y-m(1-\eta)+1)\cdots(z-m(1-\eta))} \\
&\le \left(\frac{z-m}{z-m(1-\eta)}\right)^{z-y} = \left(1+\frac{m\eta}{z-m}\right)^{y-z},
\end{aligned}
$$

and since $1+m\eta/z \ge \exp(m\eta/2z)$ and $z-y \ge (2z/m\eta)\log(m/\gamma\eta)$, we have

$$
\frac{\binom{z}{m(1-\eta)}}{\binom{y}{m(1-\eta)}}\frac{\binom{y}{m}}{\binom{z}{m}} \le \frac{\gamma\eta}{m}.
$$

Therefore

$$
\begin{aligned}
\big(\mathbb{E}|D_z|\big)^{-1} &\le \frac{1}{m(1-\eta)} + \frac{\sum_{k=0}^{m(1-\eta)}\binom{z}{k}t_k}{\sum_{k=0}^{m(1-\eta)}\binom{y}{k}t_k}\frac{\sum_{k=0}^{y}\binom{y}{k}t_k}{\sum_{k=m}^{y}\binom{y}{k}kt_k}\frac{\sum_{k=m}^{y}\binom{y}{k}kt_k}{\sum_{k=m}^{z}\binom{z}{k}kt_k} \\
&\le \frac{1}{m(1-\eta)} + \frac{\binom{z}{m(1-\eta)}}{\binom{y}{m(1-\eta)}}\frac{1}{\gamma}\frac{\binom{y}{m}}{\binom{z}{m}} \\
&\le \frac{1}{m(1-\eta)} + \frac{1}{\gamma}\frac{\gamma\eta}{m} \le \frac{1}{m(1-2\eta)}.
\end{aligned}
$$

□

We say an element $w \in \mathbb{N}$ *hits* a set $K \subseteq [w-1]$ if w selects an element in K, and we say w *misses* a set $K \subseteq [w-1]$ if w does not select any element in K.

Lemma 5.3. *Let w, z be positive integers, and $0 < \delta < 1$ a constant, satisfying*

$$
w \le z\left(1 - \frac{6\log(24\mathbb{E}|D_z|/\delta\mathbb{E}|D_w|)}{\mathbb{E}|D_z|}\right).
$$

Then, for any subset K of $[w-1]$ with $|K| = k$, where $k \le \delta w/2\mathbb{E}|D_z|$, we have

$$
\mathbb{P}(w \text{ misses } K) \le 1 - \frac{k}{w}(1-\delta)\mathbb{E}|D_w|.
$$

Note the rather curious role of z in the statement of this lemma. The point is that, under 'normal' circumstances, the probability that w hits K should be about $k\mathbb{E}|D_w|/w$, provided this quantity is not too large. However, this may fail if $|D_w|$ is typically far from its mean: if this happens, then any z satisfying the given bound will have $\mathbb{E}|D_z|$ rather larger than $\mathbb{E}|D_w|$, so the bound on k will be correspondingly more demanding.

Proof. Firstly we apply Lemma 5.2 with $m = 2\mathbb{E}|D_z|$, $\eta = 1/6$, and $\gamma = \delta\mathbb{E}|D_w|/2$. We have $\mathbb{E}|D_z| < m(1-2\eta)$ and we also have

$$
w \le z\left(1 - \frac{6\log(24\mathbb{E}|D_z|/\delta\mathbb{E}|D_w|)}{\mathbb{E}|D_z|}\right) = z\left(1 - \frac{2\log(m/\gamma\eta)}{m\eta}\right).
$$

So, to avoid a contradiction, we must have

$$
\sum_{l=m}^{w}\binom{w}{l}lt_l < \gamma\sum_{l=0}^{w}\binom{w}{l}t_l.
$$

Since $k \le \delta w/m$, we have that, for all $l < m$,

$$\left(1 - \frac{k}{w}\right)^l \le 1 - \frac{kl}{w} + \frac{(kl)^2}{2w^2} \le 1 - \frac{kl}{w}(1-\delta/2).$$

Therefore, since $\mathbb{P}(w \text{ misses } K \mid |D_w| = l) \le (1-k/w)^l$, we have

$$\begin{aligned}
\mathbb{P}(w \text{ misses } K) &\le \sum_{l=0}^{w}\left(1-\frac{k}{w}\right)^l \mathbb{P}(|D_w| = l) \\
&\le \sum_{l=0}^{m-1}\left(1-\frac{k}{w}\right)^l \mathbb{P}(|D_w| = l) + \mathbb{P}(|D_w| \ge m) \\
&\le \sum_{l=0}^{m-1}\left(1-\frac{kl}{w}(1-\delta/2)\right)\mathbb{P}(|D_w| = l) + \mathbb{P}(|D_w| \ge m) \\
&= 1 - \frac{k}{w}(1-\delta/2)\frac{\sum_{l=0}^{m-1}\binom{w}{l} l t_l}{\sum_{l=0}^{w}\binom{w}{l} t_l} \\
&= 1 - \frac{k}{w}(1-\delta/2)\left(\mathbb{E}|D_w| - \frac{\sum_{l=m}^{w}\binom{w}{l} l t_l}{\sum_{l=0}^{w}\binom{w}{l} t_l}\right) \\
&\le 1 - \frac{k}{w}(1-\delta/2)(\mathbb{E}|D_w| - \gamma) \\
&\le 1 - \frac{k}{w}(1-\delta)\mathbb{E}|D_w|.
\end{aligned}$$

where the final inequality follows from the choice of γ. □

The following results give descriptions of the large-scale structure of classical sequential growth models and the random partial orders they produce. They follow the same template: for any constant $\varepsilon > 0$, if we look on a large enough scale (take n sufficiently large) then, barring perhaps the first εn elements, the model $\mathcal{P}(\mathbf{t})$ up to stage n, (or the random partial order $P_n(\mathbf{t})$ produced by this model) has a particular structure. Importantly, the scale at which we look at depends only on the fraction ε of points we choose to ignore, and not on the particular model $\mathcal{P}(\mathbf{t})$.

Lemma 5.4. *For $\varepsilon > 0$ there exists $n_0 = n_0(\varepsilon)$ such that for all $n \ge n_0$, and any $\mathcal{P}(\mathbf{t})$, if $\mathbb{E}|D_y| \ge M$ for some $y \in [\varepsilon n, n-1-n(\log n)^{1/4}/M]$ and $(\log n)^{1/2} < M < n^{1/2}$, then $\mathbb{P}(|D_{y+n(\log n)^{1/4}/M}| \le M/(\log n)^{1/4}) < \varepsilon/8$.*

Proof. Fix $\varepsilon > 0$. Suppose that $\mathbb{E}|D_y| \ge M$, with $(\log n)^{1/2} < M < n^{1/2}$. 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

$$\sum_{i=0}^{y} i\binom{y}{i}t_i \ge M\sum_{i=0}^{y}\binom{y}{i}t_i. \tag{3}$$

Let $\alpha = (\log n)^{1/4}$ and $j = n\alpha/M$. We need to bound from above the probability

$$\mathbb{P}(|D_{y+j}| \le M/\alpha) = \frac{\sum_{i=0}^{M/\alpha}\binom{y+j}{i}t_i}{\sum_{i=0}^{y+j}\binom{y+j}{i}t_i}. \tag{4}$$

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 $j = n\alpha/M$ and $y \geq \varepsilon n$, so $j/y \leq \alpha/\varepsilon M \leq 1/\varepsilon(\log n)^{1/4}$ so for any $\eta > 0$ we have

$$\frac{\binom{y+j}{i}}{\binom{y}{i}} \geq (1-\eta)e^{ij/y} \tag{5}$$

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 $i \leq M/\alpha < n^{1/2}/(\log n)^{1/4}$ we have

$$\frac{\binom{y+j}{i}}{\binom{y}{i}} \leq e^{2Mj/y\alpha} \tag{6}$$

for sufficiently large n.

So, using the upper bound for $\binom{y+j}{i}/\binom{y}{i}$ in the numerator in (4) and the lower bound in the denominator, we have

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

and using (3) we have

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

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

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

which, after substituting $j = n\alpha/M$, $\varepsilon n \leq y \leq n$ and $\alpha = (\log n)^{1/4}$, gives

$$\mathbb{P}(|D_{y+n(\log n)^{1/4}/M}| \leq M/(\log n)^{1/4}) \leq \frac{e^{2/\varepsilon-1}}{(1-\eta)(\log n)^{1/4}}$$

for sufficiently large n. Therefore

$$\mathbb{P}(|D_{y+n(\log n)^{1/4}/M}| \leq M/(\log n)^{1/4}) < \varepsilon/8$$

for sufficiently large n, as required. □

For $w \in \mathbb{N}$, $w \neq 0, 1$, define $\rho_w = \mathbb{E}|D_w|/w\log w$ and define the distance between $y, z \in \mathbb{N}$, for $0 < y < z$, to be

$$d(y,z) = \sum_{w=y+1}^{z}\rho_w. \tag{7}$$

We now state our main result, which says that the distance function reliably predicts comparability of elements in $P_n(\mathbf{t})$.

Theorem 5.5. *For $\varepsilon > 0$, there exists $n_0 = n_0(\varepsilon)$ such that for all $n \geq n_0$, and any $\mathcal{P}(\mathbf{t})$, the following holds. For $y, z \in [\varepsilon n, n-1]$ with $y < z$, provided that $\mathbb{E}|D_z| < (\log n)^{5/4}$, we have:*

(i) if $d(y,z) < 1-\varepsilon$, then $\mathbb{P}(y < z \text{ in } P_n(\mathbf{t})) < \varepsilon$,

(ii) if $d(y,z) > 1+\varepsilon$, then $\mathbb{P}(y < z \text{ in } P_n(\mathbf{t})) > 1-\varepsilon$.

Proof. We prove (i) by a path-counting argument, similar to several arguments used by Georgiou [9].

We write $E(y,z)$ for the expected number of sequences $y = r_0, r_1, \ldots, r_k = z$, such that r_i selects r_{i-1} (i.e., $r_{i-1} \in D_{r_i}$) for each $i = 1, \ldots, k$. Note that $y < z$ in $P_n(\mathbf{t})$ if and only if there is such a sequence, for some $k \geq 1$. Next, we note that, for elements $r > s$, the probability that r selects s, is exactly $\mathbb{E}|D_r|/r$, so in particular is independent of s. This yields a formula for $E(y,z)$:

$$\begin{aligned} E(y,z) = \left[\prod_{r=y+1}^{z-1} \big(1 + \mathbb{P}(r \text{ selects } s)\big)\right] \mathbb{P}(z \text{ selects } s) &= \left[\prod_{r=y+1}^{z-1} \left(1 + \frac{\mathbb{E}|D_r|}{r}\right)\right] \frac{\mathbb{E}|D_z|}{z} \\ &\leq \frac{\mathbb{E}|D_z|}{z} \exp\left\{\sum_{r=y+1}^{z-1} \frac{\mathbb{E}|D_r|}{r}\right\} \\ &= \frac{\mathbb{E}|D_z|}{z} \exp\left\{\sum_{r=y+1}^{z-1} \rho_r \log r\right\} \\ &\leq \frac{\mathbb{E}|D_z|}{z} \exp\left\{\log z \, d(y, z-1)\right\} \\ &\leq \frac{\mathbb{E}|D_z|}{z} z^{1-\varepsilon} = \frac{\mathbb{E}|D_z|}{z^{\varepsilon}} < \varepsilon, \end{aligned}$$

for sufficiently large n. So, by Markov's inequality, we have $\mathbb{P}(y < z \text{ in } P_n(\mathbf{t})) < E(y,z) < \varepsilon$ as required.

To prove (ii), we show that with high probability, when we reach a distance $1+\varepsilon/2$ from y, the up-set of y has grown to a reasonable size. We then show that this size increases in a predictable manner so that almost all elements at distance greater than $1+\varepsilon$ from y are in the up-set of y.

For now, we assume that $\mathbb{E}|D_y| > 2(\log n)^{3/4}$; however, we will show later that (ii) also holds when this condition is not satisfied.

For $i = 0, \ldots, 3$, let $w_i = \min\{w : d(y,w) > 1 + \frac{i+4}{8}\varepsilon\}$. Note that $d(y, w_0) \approx 1 + \varepsilon/2$, $d(w_i, w_{i+1}) \approx \varepsilon/8$, for $i = 0, \ldots, 2$ and $d(w_3, z) > \varepsilon/8$. Let $\sigma = \mathbb{E}|D_{w_3}|/y \log y$. So, $\rho_w = \mathbb{E}|D_w|/w \log w \leq \mathbb{E}|D_{w_3}|/y \log y = \sigma$ for all $w \in [y, w_3]$. Set $k_0 = y/2\mathbb{E}|D_{w_3}| \log y$.

We want to apply Lemma 5.3, for $w \in [y, w_2 - 1]$, with $z = w_3$, $\delta = (\log y)^{-1}$, to a set K of size $k \leq k_0$. Since $k \leq k_0 \leq \delta w / 2\mathbb{E}|D_{w_3}|$ for all $w \in [y, w_2 - 1]$, it remains to show that

$$w \leq w_3 \left(1 - \frac{6 \log(24\mathbb{E}|D_{w_3}|/\delta \mathbb{E}|D_w|)}{\mathbb{E}|D_{w_3}|}\right).$$

We use the fact that $d(w, w_3) > \varepsilon/8$ for all $w \in [y, w_2 - 1]$. Since

$$d(w, w_3) = \sum_{u=w+1}^{w_3} \rho_w < \frac{\mathbb{E}|D_{w_3}|}{w \log w}(w_3 - w)$$

we have $w_3 - w > \varepsilon w \log w / 8\mathbb{E}|D_{w_3}|$, so it is enough to show that

$$\frac{\varepsilon w \log w}{8\mathbb{E}|D_{w_3}|} > \frac{6w_3 \log(24\mathbb{E}|D_{w_3}|/\delta\mathbb{E}|D_w|)}{\mathbb{E}|D_{w_3}|}. \tag{8}$$

Now note that $w, w_3 \in [\varepsilon n, n]$, and that $\mathbb{E}|D_{w_3}| < (\log n)^{5/4}$, $\mathbb{E}|D_w| > 2(\log n)^{3/4}$. Therefore

$$\frac{\varepsilon w \log w}{8} > \frac{\varepsilon^2 n \log(\varepsilon n)}{8}$$

and

$$6w_3 \log(24\mathbb{E}|D_{w_3}|/\delta\mathbb{E}|D_w|) < 6n \log(12(\log n)^{1/2} \log y) < 6n \log(12(\log n)^{3/2}).$$

Since

$$\frac{\varepsilon^2 n \log(\varepsilon n)}{8} > 6n \log(12(\log n)^{3/2})$$

holds for sufficiently large n, we have inequality (8) as required, for sufficiently large n. Therefore, using Lemma 5.3,

$$\begin{aligned} \mathbb{P}(w \text{ misses a set } K \text{ of size } k) &\le 1 - \frac{k}{w}\mathbb{E}|D_w|(1 - (\log y)^{-1}) \\ &= 1 - k\rho_w \log w(1 - (\log y)^{-1}), \end{aligned}$$

for all $w \in [y, w_2 - 1]$ and $k \le k_0$.

Set $A_k = k(\log y - 1)$, so that for all $w \in [y, w_2 - 1]$ and $k \le k_0$,

$$\mathbb{P}(w \text{ misses a set } K \text{ of size } k) \le 1 - \rho_w A_k.$$

Set $u_k = \min\{u : |U[y] \cap [u]| \ge k\}$. Note that $u_1 = y$. Define the *waiting distance*, $W_{k+1} = d(u_k, u_{k+1})$. Then, for $u > u_k$,

$$\begin{aligned} \mathbb{P}(W_{k+1} > \rho_{u_k+1} + \ldots + \rho_u) &= \prod_{j=u_k+1}^{u} \mathbb{P}(j \text{ misses } U[y] \cap [u_k]) \\ &\le \prod_{j=u_k+1}^{u} (1 - \rho_j A_k) \\ &\le \exp\left(-A_k \sum_{j=u_k+1}^{u} \rho_j\right). \end{aligned}$$

For any $x \in \mathbb{R}_+$, set u^* to be the maximum u such that $\rho_{u_k+1} + \ldots + \rho_{u^*} \le x + \sigma$. We have

$$\begin{aligned} \mathbb{P}(W_{k+1} - \sigma > x) = \mathbb{P}(W_{k+1} > x + \sigma) &\le \mathbb{P}(W_{k+1} > \rho_{u_k+1} + \ldots + \rho_{u^*}) \\ &\le \exp\left(-A_k \sum_{j=u_k+1}^{u^*} \rho_j\right) \le \exp(-A_k x), \end{aligned}$$

since $\rho_{u_k+1} + \ldots + \rho_{u^*} + \rho_{u^*+1} > x + \sigma$ and $\rho_{u^*+1} \le \sigma$, so $\rho_{u_k+1} + \ldots + \rho_{u^*} > x$. So, for all $k \le k_0$, and all $x \in \mathbb{R}_+$,

$$\mathbb{P}(W_{k+1} - \sigma > x) \le e^{-A_k x},$$

independent of the history up to u_k (provided we never get past w_2).

Now define $(X_{k+1})_{k=1}^{k_0}$ to be independent random variables, with

$$\mathbb{P}(X_{k+1} > x) = e^{-A_k x}$$

for all x. So,

$$\mathbb{P}\left(\sum_{k=1}^{k_0}(W_{k+1}-\sigma) > M\right) \le \mathbb{P}\left(\sum_{k=1}^{k_0} X_{k+1} > M\right).$$

But each X_{k+1} is an exponential random variable with mean $1/A_k$ and variance $1/A_k^2$, so $\sum_{k=1}^{k_0} X_{k+1}$ has mean

$$\sum_{k=1}^{k_0}\frac{1}{A_k} = \frac{1}{\log y - 1}\sum_{k=1}^{k_0}\frac{1}{k} \le \frac{\log y - \log(m\log y)}{\log y - 1} \le 1,$$

and variance

$$\sum_{k=1}^{k_0}\frac{1}{A_k^2} = \frac{1}{(\log y-1)^2}\sum_{k=1}^{k_0}\frac{1}{k^2} \le \frac{2}{\log^2 y}$$

So, by Chebyshev's inequality,

$$\mathbb{P}\left(\sum_{k=1}^{k_0} X_{k+1} \ge 1 + \frac{2}{\varepsilon\log y}\right) < \varepsilon^2/2 < \varepsilon/2,$$

for $\varepsilon < 1$. So,

$$\mathbb{P}\left(\sum_{k=1}^{k_0} W_{k+1} \ge \sigma k_0 + 1 + \frac{2}{\varepsilon\log y}\right) < \varepsilon/2,$$

and since $\sigma k_0 = 1/2\log^2 y$, we have that with probability at least $1-\varepsilon/2$, the total waiting distance $\sum_{k=1}^{k_0} W_{k+1}$ is less than $1 + O(1/\log y) < 1+\varepsilon/2$. This means that for w_0 with $d(y, w_0) > 1+\varepsilon/2$, the size of the up-set $|U[y]\cap[w_0]|$ is greater than $k_0 = y/2\mathbb{E}|D_{w_3}|\log y$ with probability at least $1-\varepsilon/2$.

We now show that the size of the up-set $U[y]\cap[w]$ increases in a predictable manner as w increases from w_0 to w_2, by repeated application of the following claim.

Claim. *For $i = 0, 1$, if $K_i \subset U[y]\cap[w_i]$ with $|K_i| < y/2\mathbb{E}|D_{w_3}|\log\log y$, then*

$$\mathbb{P}(|U[y]\cap[w_{i+1}]| < |K_i|(\log y)^{3/4}) < \varepsilon/8.$$

Proof of Claim. We apply Lemma 5.3 for $w \in [w_i, w_{i+1}-1]$, $z = w_3$ with $\delta = 1/\log\log y$. Since $K_i < y/2\mathbb{E}|D_{w_3}|\log\log y < \delta w/2\mathbb{E}|D_{w_3}|$ for all $w\in[w_i, w_{i+1}]$, and as before

$$w \le w_3\left(1 - \frac{6\log(24\mathbb{E}|D_{w_3}|/\delta\mathbb{E}|D_w|)}{\mathbb{E}|D_{w_3}|}\right),$$

we have

$$\mathbb{P}(w \text{ misses } K_i) \le 1 - \frac{|K_i|}{w}(1-\delta)\mathbb{E}|D_w|.$$

So, for each $w\in[w_i, w_{i+1}]$,

$$\mathbb{P}(w \text{ hits } K_i) \ge \frac{|K_i|}{w}(1-\delta)\mathbb{E}|D_w| \ge |K_i|(1-\delta)\rho_w\log y.$$

and therefore

$$\mathbb{E}(\text{number of hits}) \ge |K_i|(1-\delta)\log y\sum_{w=w_i}^{w_{i+1}}\rho_w = (\varepsilon/8)|K_i|(1-\delta)\log y.$$

Since the number of hits is a sum of independent 0-1 Bernoulli trials we have

$$\mathbb{V}\mathrm{ar}(\text{number of hits}) = \sum_w \mathbb{P}(w \text{ hits } K_i)\mathbb{P}(w \text{ misses } K_i) \le \mathbb{E}(\text{number of hits}),$$

so, for sufficiently large n, with probability greater than $1-\varepsilon/8$ the number of hits is at least $|K_i|(\log y)^{3/4}$. Since each of these hits is an element placed above $K_i \subset U[y] \cap [w_i]$, we have a set of at least $|K_i|(\log y)^{3/4}$ elements in $U[y] \cap [w_{i+1}]$, with probability at least $1-\varepsilon/8$. □

We apply the claim twice: let K_0 be a subset of $U[y] \cap [w_0]$ of size $y/2\mathbb{E}|D_{w_3}|\log y$, so $U[y] \cap [w_1]$ is of size at least $y/2\mathbb{E}|D_{w_3}|(\log y)^{1/4}$; let K_1 be a subset of $U[y] \cap [w_1]$ of size $y/2\mathbb{E}|D_{w_3}|(\log y)^{1/4}$, so $U[y] \cap [w_2]$ is of size at least $y(\log y)^{1/2}/2\mathbb{E}|D_{w_3}|$.

Finally, we use Lemma 5.4, to show that $|D_z|$ is sufficiently large so that the probability that z misses $U[y]\cap[w_2]$ is small. We need to apply the lemma to $y = w_3$. We have $(\log n)^{1/2} < \mathbb{E}|D_{w_3}| < n^{1/2}$ so it remains to check that $w_3 + n(\log n)^{1/4}/\mathbb{E}|D_{w_3}| < z$. We use the fact that $d(w_3, z) > \varepsilon/8$:

$$d(w_3, z) \le \frac{\mathbb{E}|D_z|}{w_3 \log w_3}(z - w_3) < \frac{(\log n)^{5/4}}{\varepsilon n \log(\varepsilon n)}(z - w_3),$$

so $z - w_3 \ge \varepsilon^2 n \log(\varepsilon n)/8(\log n)^{5/4}$. Since $\mathbb{E}|D_{w_3}| > 2(\log n)^{3/4}$ we have

$$z - w_3 > \varepsilon^2 n \log(\varepsilon n)/4\mathbb{E}|D_{w_3}|(\log n)^{1/2} > n(\log n)^{1/4}/\mathbb{E}|D_{w_3}|,$$

as required.

So, by Lemma 5.4, we have $\mathbb{P}(|D_z| \le \mathbb{E}|D_{w_3}|/(\log n)^{1/4}) < \varepsilon/8$. Provided that $|D_z|$ is at least $\mathbb{E}|D_{w_3}|/(\log n)^{1/4}$, then

$$\begin{aligned}\mathbb{P}(z \text{ misses } U[y] \cap [w_2]) &\le \left(1 - \frac{y(\log y)^{1/2}}{2\mathbb{E}|D_{w_3}|z}\right)^{\mathbb{E}|D_{w_3}|/(\log n)^{1/4}} \\ &\le \exp\left(-\frac{y(\log y)^{1/2}}{2z(\log n)^{1/4}}\right) \\ &\le \exp\left(-\frac{\varepsilon(\log(\varepsilon n))^{1/2}}{2(\log n)^{1/4}}\right) \le \varepsilon/8\end{aligned}$$

for sufficiently large n.

Combining these results, we have that $\mathbb{P}(y < z \text{ in } P_n(\mathbf{t})) > 1 - \varepsilon$, as required.

Recall that we assumed that $\mathbb{E}|D_y| > 2(\log n)^{3/4}$. To complete the proof, we now show that (ii) also holds when $\mathbb{E}|D_y| \le 2(\log n)^{3/4}$. Since $d(y, z) > 1 + \varepsilon$, and

$$d(y, z) < \frac{\mathbb{E}|D_z|}{y \log y}(z - y) < \frac{\mathbb{E}|D_z|}{\varepsilon \log(\varepsilon n)}$$

we have $\mathbb{E}|D_z| > \varepsilon \log(\varepsilon n) > 2(\log n)^{3/4}$ for sufficiently large n. Since $\mathbb{E}|D_w|$ is increasing, there exists y' with $y < y' < z$ such that $\mathbb{E}|D_{y'}| > 2(\log n)^{3/4}$ and $\mathbb{E}|D_w| \le 2(\log n)^{3/4}$ for all $w \in [y, y'-1]$.

So,

$$d(y, y'-1) < \frac{\mathbb{E}|D_{y'-1}|}{y \log y}(y' - y) < \frac{(\log n)^{3/4}}{\varepsilon \log(\varepsilon n)} < \varepsilon/4$$

for sufficiently large n. Moreover, since $\mathbb{E}|D_z| < (\log n)^{5/4}$,

$$d(y'-1, y'+n/\sqrt{\log n}) < \frac{(\log n)^{5/4}}{y' \log y'}(n/\sqrt{\log n}+1) < \frac{2(\log n)^{3/4}}{\varepsilon \log(\varepsilon n)} < \varepsilon/4$$

for sufficiently large n. Therefore, for all $w \in [y', y'+n/\sqrt{\log n}]$, we have $d(w, z) > 1+\varepsilon/2$, and $\mathbb{E}|D_w| \geq 2(\log n)^{3/4}$. So, we can apply the above argument (rather, a minor modification, replacing ε by $\varepsilon/2$), which shows that

$$\mathbb{P}(w < z \text{ in } P_n(\mathbf{t})) > 1-\varepsilon/2,$$

for all $w \in [y', y'+n/\sqrt{\log n}]$. Note that these events are not independent.

Since $\mathbb{E}|D_w| > 2(\log n)^{3/4}$,

$$\mathbb{P}(w \text{ selects } y) = \frac{\mathbb{E}|D_w|}{w} > \frac{(\log n)^{3/4}}{n},$$

for all $w \in [y', y'+n/\sqrt{\log n}]$, so the expected number of elements w that select y is at least $(\log n)^{1/4}$. Since the variance is bounded by the expectation, we have that, with probability at least $1-\varepsilon/2$, there exists at least one element in $[y', y'+n/\sqrt{\log n}]$ that selects y, for sufficiently large n.

So, let u be the first element in $[y', y'+n/\sqrt{\log n}]$ to select y, with probability at least $1-\varepsilon/2$ there is such an element. Then, with probability at least $1-\varepsilon/2$ the element z is above u in $P_n(\mathbf{t})$. Since u is above y, we have that z is also above y in $P_n(\mathbf{t})$ as required. □

We have shown that for any $\mathcal{P}(\mathbf{t})$ we can define a distance function that, for any $\varepsilon > 0$ and sufficiently large n, reliably predicts the comparability of elements in the set $[\varepsilon n, n-1]$. We have relied on the condition $\mathbb{E}|D_z| < (\log n)^{5/4}$, which at first sight appears rather arbitrary. However, an upper bound is necessary: when ρ_w can suddenly increase, the distance $d(y, z)$ may become artificially large so that the function wrongly predicts y, z to be comparable. It is interesting to note that a similar lower bound on $\mathbb{E}|D_w|$ is not needed to guarantee the correct prediction of incomparability: note in the proof of Theorem 5.5, that the temporary condition $\mathbb{E}|D_y| > 2(\log n)^{3/4}$ could be removed, essentially because $d(y, z)$ is always less than 1 when $\mathbb{E}|D_z| = o(\log n)$.

We now show that for $\mathbb{E}|D_w| > (\log n)^{5/4}$ almost all pairs in $[w, n-1]$ are comparable.

Theorem 5.6. *For $\varepsilon > 0$ there exists $n_0 = n_0(\varepsilon)$ such that for all $n \geq n_0$ and any $\mathcal{P}(\mathbf{t})$, the following is true. If there exists $z_C \in [\varepsilon n, (1-\varepsilon)n]$ with $\mathbb{E}|D_{z_C}| \geq (\log n)^{5/4}$, then, for $y \in [\varepsilon n, n]$, $z \in [z_C + \varepsilon n, n-1]$ with $z - y > \varepsilon n$,*

$$\mathbb{P}(y < z \textit{ in } P_n(\mathbf{t})) > 1-\varepsilon.$$

We omit the proof as it is very similar to that of part (ii) of Theorem 5.5.

To summarise, we have the following large-scale description of $P(\mathbf{t})$. For $\varepsilon > 0$ and large n, we can partition the set $[n-1]$ into three regions. There is a small 'unknown' set $[0, \varepsilon n]$ which we do not describe. The remainder is partitioned at the point z_C, where $\mathbb{E}|D_{z_C}| = (\log n)^{5/4}$. The set $[\varepsilon n, z_C]$ looks very similar to a random graph order, in that near elements are likely to be incomparable and far apart elements likely to be comparable, as predicted by the distance function $d(y, z)$. Unlike the random graph order, this critical distance can vary over the length of the set, depending on how $\mathbb{E}|D_w|$ increases. The set $[z_C, n-1]$ is typically very dense, so that all pairs of elements are likely to be comparable. Note that either region could be empty, and any pair of elements with $y \in [\varepsilon n, z_C]$ and $z \in [z_C, n-1]$ is likely to be comparable.

As an example, consider a random graph order $P_{n,p}$ with $\lim_{n\to\infty} p^{-1}\log p^{-1}/n = c$ for some $0 < c < 1$. For $\varepsilon > 0$, since $p \sim \log n/(cn)$, and $\mathbb{E}|D_y| = py$, we have $\mathbb{E}|D_y| \sim y\log n/(cn) \leq c^{-1}\log n$ for all $y \in [\varepsilon n, n-1]$. This means that we take $z_C = n-1$ and for any two points $y < z \in [\varepsilon n, n-1]$, whether they are comparable depends on the value of

$$d(y,z) = \sum_{w=y+1}^{z} \frac{\mathbb{E}|D_w|}{w\log w} = \sum_{w=y+1}^{z} \frac{p}{\log w} \sim \frac{z-y}{cn}.$$

So for large enough n, if $z-y > (1+\varepsilon)cn$ then z is likely to be ordered above y in $P_{n,p}$, whereas if $|z-y| < (1-\varepsilon)cn$ then y and z are likely to be incomparable in $P_{n,p}$. Note that this agrees with the continuum limit as given by Theorem 4.2.

We can now prove Theorem 1.5.

Proof of Theorem 1.5. In fact, we prove the equivalent statement that for all ε with $0 < \varepsilon < 1/4$ there exists an n_1 such that for any classical sequential growth model $\mathcal{P}(\mathbf{t})$ and for all $n \geq n_1$, there exists a semiorder S on $\{0,\dots,n-1\}$ such that $\mathbb{E}\Delta(P_n(\mathbf{t}),S) < 6\sqrt{\varepsilon}n^2$. So, fix ε with $0 < \varepsilon < 1/4$, and let n be sufficiently large that both Theorems 5.5 and 5.6 hold. For any $\mathcal{P}(\mathbf{t})$ we define the following order $\prec_S$ on $[n-1]$. If there exists a $z \in [\varepsilon n, (1-\varepsilon)n]$ with $\mathbb{E}|D_z| \geq (\log n)^{5/4}$, then define z_C as the minimum such z; otherwise set $z_C = (1-\varepsilon)n$. For all $y < z$ with $z \geq z_C$ set $y \prec_S z$. For all $y < z$ with $z < z_C$, set $y \prec_S z$ if and only if $d(y,z) > 1$. Let $P_n(\mathbf{t})$ be a random partial order on $[n-1]$ produced according to $\mathcal{P}(\mathbf{t})$. We will show that the expected number of pairs $(y,z) \in [n-1]^{(2)}$ on which $P_n(\mathbf{t})$ and $\prec_S$ differ is less than $6\sqrt{\varepsilon}n^2$. We can assume that all pairs $y < z$ with $y < \varepsilon n$, or $z_C < z < z_C + \varepsilon n$ contribute to this difference, a total of no more than $2\varepsilon n^2$ pairs. For all other pairs, either (a) $\varepsilon n \leq z < z_C$, or (b) $z \geq z_C + \varepsilon n$. For pairs satisfying (a), the orders agree with probability $1-\varepsilon$ when $|d(y,z)-1| > \varepsilon$ (using Theorem 5.5), and so we need to show that the number of pairs with $|d(y,z)-1| \leq \varepsilon$ is small. For $y \in [\varepsilon n, z_C]$ let z_0 be the smallest z such that $d(y,z) \geq 1-\varepsilon$, and define z_1 as the largest z such that $d(y,z) \leq 1+\varepsilon$. So,

$$\sum_{w=y+1}^{z_0} \frac{\mathbb{E}|D_w|}{w\log w} \geq 1-\varepsilon, \quad \text{and} \quad \sum_{w=z_0+1}^{z_1} \frac{\mathbb{E}|D_w|}{w\log w} \leq 2\varepsilon.$$

Taking the first inequality and bounding the sum from above gives $1-\varepsilon \leq \mathbb{E}|D_{z_0}|n/(y\log\varepsilon n)$, and bounding the sum in the second inequality from below gives $2\varepsilon \geq (z_1-z_0)\mathbb{E}|D_{z_0}|/(z_1\log n)$. Combining these gives

$$z_1 - z_0 \leq \frac{2\varepsilon n z_1 \log n}{(1-\varepsilon)y\log\varepsilon n} \leq 4\varepsilon\frac{n^2}{y}$$

for sufficiently large n. So, for each $y \in [\varepsilon n, n-1]$ there are at most $4\varepsilon n^2/y$ pairs with $|d(y,z)-1| \leq \varepsilon$. Therefore the total number of such pairs is at most

$$\sum_{y=\varepsilon n}^{n} \frac{4\varepsilon n^2}{y} \leq 4\varepsilon n^2 \int_{\varepsilon n-1}^{n} \frac{dx}{x} \leq 4\varepsilon n^2\log(1/\varepsilon) \leq 4\sqrt{\varepsilon}n^2.$$

Finally, for pairs satisfying (b), we know that if $z-y > \varepsilon n$ then Theorem 5.6 implies that the orders agree with probability $1-\varepsilon$, and there are at most εn^2 pairs with $z-y \leq \varepsilon n$.

Therefore the total expected number of pairs on which the orders disagree is at most $2\varepsilon n^2 + 4\sqrt{\varepsilon}n^2 + \varepsilon n^2 + \varepsilon n^2 = 4\sqrt{\varepsilon}n^2 + 4\varepsilon n^2 < 6\sqrt{\varepsilon}n^2$, since $\varepsilon < 1/4$. □

As mentioned earlier, the order of the quantifiers in Theorem 1.5 is significant: the random poset on $[n-1]$ produced by any classical sequential growth model is arbitrarily close to being a semiorder, for large enough n. This means we are able to prove that any continuum limit of a sequence of classical sequential growth models is also arbitrarily close to being a semiorder, hence it is an almost-semiorder, as claimed by Theorem 1.6.

Proof of Theorem 1.6. Suppose P_∞ is the continuum limit of $(\mathcal{P}_n)_{n=1}^\infty$. Recall that H and L are the four-element partial orders in Figure 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 $\lambda(L;P_\infty) = 0$.

For P and Q finite partial orders on the same ground set X, and $k \geq 2$, define $\Delta_k(P,Q)$ to be the number of subsets of X of size k on which P and Q differ. That is,

$$\Delta_k(P,Q) = \#\{A \in X^{(k)} : P \text{ induces a different partial order from } Q \text{ on } A\}$$

Note that Δ_2 is identical to Δ as defined in Section 1. For fixed P, Q, X and all $k \leq l$, we have that every subset of X of size k is contained in less than $\binom{|X|}{l-k} \leq |X|^{l-k}/(l-k)!$ subsets of X of size l, and therefore

$$\Delta_l(P,Q) \leq \frac{|X|^{l-k}}{(l-k)!}\Delta_k(P,Q) \tag{9}$$

for $2 \leq k \leq l$.

Now, fix $\varepsilon > 0$ and apply Theorem 1.5. So, for each n greater than the n_0 given by Theorem 1.5, there exists a semiorder S, such that $\mathbb{E}\Delta_2(P_n,S) \leq \varepsilon n^2$, where P_n is the random partial order produced according to $\mathcal{P}_n$. Therefore, using equation (9), we have $\mathbb{E}\Delta_4(P_n,S) \leq \varepsilon n^4/2$. Consider a sample of four points $\{x_1,x_2,x_3,x_4\}$ from P_n. Unless the sample is equal to one of the at most $\varepsilon n^4/2$ sets on which P_n and S differ then the order on $\{x_1,x_2,x_3,x_4\}$ is not equal to H or L. (This is because of the definition of S as a semiorder.) Therefore, we have that both $\mathbb{E}\lambda(H;P_n)$ and $\mathbb{E}\lambda(L;P_n)$ are less than 12ε. □

6 Existence of continuum limits

Theorem 1.5 tells us about the global structure of a particular classical sequential growth model in the same way that Theorems 3.2–3.4 of Pittel and Tungol tell us about the structure of a particular random graph order. Indeed, as the results of Pittel and Tungol could be applied to prove results about continuum limits of sequences of random graph orders, we now apply Theorem 1.5 to prove a result about sequences of classical sequential growth models, which extends Theorem 4.2.

Let us embellish our current notation: we will be considering sequences $(\mathcal{P}_n)_{n=1}^\infty$ of classical sequential growth models, so for each n, the model $\mathcal{P}_n$ is described by a sequence $\mathbf{t}^{(n)}$, and we define

$$\rho_w^{(n)} = \frac{\mathbb{E}|D_w|}{w\log w}, \qquad d^{(n)}(y,z) = \sum_{w=y+1}^{z} \rho_w^{(n)},$$

where the expectation $\mathbb{E}|D_w|$ is defined on the model $\mathcal{P}_n$, and depends only on $\mathbf{t}^{(n)}$.

For each n, we define a step function r_n from $[0,1]$ to $\mathbb{R}$ by

$$r_n(x) = \begin{cases} n\rho_{\lceil xn\rceil}^{(n)} & \text{for } x > 1/n, \\ 0 & \text{otherwise.} \end{cases}$$

We make some simple observations about the functions r_n.

Lemma 6.1. *For $\varepsilon > 0$, there exists n_0 such that for any $x, y \in [\varepsilon, 1]$ with $x < y$,*

$$x\, r_n(x) \le (1+\varepsilon) y\, r_n(y)$$

for all $n \ge n_0$.

Proof. This follows simply from the definition of r_n. Fix $\varepsilon > 0$. Then for any $x, y \in [\varepsilon, 1]$ with $x < y$, since $\mathbb{E}|D_w|$ is increasing in w, we have

$$r_n(x) = \frac{n\mathbb{E}|D_{\lceil xn \rceil}|}{\lceil xn \rceil \log \lceil xn \rceil} \le \frac{n\mathbb{E}|D_{\lceil yn \rceil}|}{\lceil yn \rceil \log \lceil yn \rceil} \frac{\lceil yn \rceil \log \lceil yn \rceil}{\lceil xn \rceil \log \lceil xn \rceil} \le r_n(y)(1+\varepsilon)\frac{y}{x}$$

for sufficiently large n. □

Lemma 6.2. *For $\varepsilon > 0$, there exists n_0 such that for any $x, y \in [\varepsilon, 1]$ with $x < y$,*

$$-\frac{1}{n} r_n(x) \le d^{(n)}(\lceil xn \rceil, \lceil yn \rceil) - \int_x^y r_n(t)dt \le \frac{1}{n} r_n(y)$$

for all $n \ge n_0$.

Proof. We integrate $r_n(x)$ using the fact that it is a step function: for all integers k, l, with $1 \le k < l \le n$,

$$\int_{k/n}^{l/n} r_n(t)dt = \frac{1}{n} \sum_{w=k+1}^{l} r_n(w/n) = \sum_{w=k+1}^{l} \rho_w^{(n)} = d^{(n)}(k, l).$$

Now, fix $\varepsilon > 0$ and take $n_0 > 1/\varepsilon$. Then, for any $x, y \in [\varepsilon, 1]$ with $x < y$,and $n \ge n_0$,

$$\int_x^y r_n(t)dt \le \int_{(\lceil xn \rceil - 1)/n}^{\lceil yn \rceil/n} r_n(t)dt = d^{(n)}(\lceil xn \rceil - 1, \lceil yn \rceil) = \rho_{\lceil xn \rceil}^{(n)} + d^{(n)}(\lceil xn \rceil, \lceil yn \rceil)$$

and

$$\int_x^y r_n(t)dt \ge \int_{\lceil xn \rceil/n}^{(\lceil yn \rceil - 1)/n} r_n(t)dt = d^{(n)}(\lceil xn \rceil, \lceil yn \rceil - 1) = d^{(n)}(\lceil xn \rceil, \lceil yn \rceil) - \rho_{\lceil yn \rceil}^{(n)},$$

which yields the required inequality after rearrangement and the substitution $r_n(x) = n\rho_{\lceil xn \rceil}^{(n)}$. □

Lemma 6.3. *For $\varepsilon > 0$, and any interval $I = [\alpha, \beta] \subseteq [\varepsilon, 1]$, if $\limsup_{n\to\infty} r_n(\beta)$ is finite then $\limsup_{n\to\infty} r_n(x)$ is integrable on I and*

$$\limsup_{n\to\infty} \int_I r_n(t)dt \le \int_I \limsup_{n\to\infty} r_n(t)dt. \tag{10}$$

Furthermore, if $\lim_{n\to\infty} r_n(x)$ exists for almost every $x \in I$ then

$$\lim_{n\to\infty} \int_I r_n(t)dt = \int_I \lim_{n\to\infty} r_n(t)dt. \tag{11}$$

Proof. The fact the the limits pass through the integrals as claimed above follows from a simple application of the Dominated Convergence Theorem/Fatou's Lemma. We need to show that r_n is dominated by an integrable function on I. We show this using Lemma 6.1.

Fix $\varepsilon > 0$ and $I = [\alpha, \beta] \subseteq [\varepsilon, 1]$. By Lemma 6.1, there exists an n_0 such that

$$x\, r_n(x) \leq (1+\varepsilon)\beta\, r_n(\beta)$$

for all $x \in I$, all $n \geq n_0$. But $\limsup_{n\to\infty} r_n(\beta)$ is finite, therefore there exists some constant K and n_1 such that $r_n(\beta) \leq K$ for all $n \geq n_1$. Therefore, defining $\Phi(x) = (1+\varepsilon)\beta K/x$, we have $r_n(x) \leq \Phi(x)$ for all $x \in I$, all $n \geq \max\{n_0, n_1\}$. Clearly $\Phi(x)$ is integrable on I, and therefore Fatou's Lemma implies that $\limsup_{n\to\infty} r_n(x)$ is integrable on I and

$$\limsup_{n\to\infty} \int_I r_n(t)dt \leq \int_I \limsup_{n\to\infty} r_n(t)dt$$

as required.

Furthermore, since $r_n(x)$ is non-negative on I, for all n, Fatou's Lemma also implies that

$$\int_I \liminf_{n\to\infty} r_n(t)dt \leq \liminf_{n\to\infty} \int_I r_n(t)dt$$

and therefore

$$\int_I \liminf_{n\to\infty} r_n(t)dt \leq \liminf_{n\to\infty} \int_I r_n(t)dt \leq \limsup_{n\to\infty} \int_I r_n(t)dt \leq \int_I \limsup_{n\to\infty} r_n(t)dt.$$

Now, if $\lim_{n\to\infty} r_n(x)$ exists for almost every $x \in I$, then the left- and right-most terms in the above inequality are equal, and the inequality collapses to equation (11) as required. □

We are now in a position to prove Theorem 1.8. In fact we will show the following stronger result: part (i) is exactly Theorem 1.8 as stated earlier.

Theorem 6.4. *A sequence $(\mathcal{P}_n)_{n=1}^{\infty}$ of classical sequential growth models with associated functions r_n as defined above has a continuum limit when either*

(i) $r(x) = \lim_{n\to\infty} r_n(x) \in [0, \infty]$ *exists for almost every* x *in* $[0, 1]$, *or*

(ii) there exists $y_C \in [0, 1]$ *such that*

- *a)* $\lim_{n\to\infty} r_n(x) = \infty$ *for all* $x \in (y_C, 1]$,
- *b)* $r_n(x)$ *is bounded for all* $x \in [0, y_C)$, *and*
- *c)* $\displaystyle\int_{\varepsilon}^{y_C-\varepsilon} R(t)dt \leq 1$, *for all* $\varepsilon > 0$, *where* $R(x) = \limsup_{n\to\infty} r_n(x)$.

The continuum limit in each case is

(i) T_r,

(ii) $T_{\tilde{R}}$, *where* $\tilde{R}(x) = 0$ *for* $x \in [0, y_C)$, *and* $\tilde{R}(x) = \infty$ *for* $x \in (y_C, 1]$.

Proof. Note the similarity in style to Theorem 4.2; the proof will also follow a similar style. Let us quickly recall the proof of Theorem 4.2, where there were three different cases. For the trivial case where the continuum limit was an antichain, it was enough to show that the probability of a sample of two elements from the finite partial order P_n being a chain tended to zero as n tended to infinity. Similarly, at the other extreme, for the case of the continuum limit being a chain, we showed that the probability of a sample being a 2-element antichain tended to zero. For the non-trivial case, we defined an intermediate partially ordered measure space $\bar{P}_n$ on $[0,1]$, for each n, in such a way that $\lim_{n\to\infty} \lambda(Q; P_n) = \lim_{n\to\infty} \lambda(Q; \bar{P}_n)$ for all finite partial orders Q. We then showed that for any sample S from $[0,1]$, the probability of the order on S induced by $\bar{P}_n$ differing from that induced by the claimed continuum limit tended to zero as n tended to infinity.

Here, we first prove case (ii), which roughly corresponds to a combination of the trivial cases (i) and (iii) of Theorem 4.2.

We need to show that the continuum limit is $T_{\tilde{R}} = ([0,1], \mathcal{B}, \mu_L, \prec)$. The definition of $\tilde{R}$ implies that $x \prec y$ if and only if $y > y_C$ and $x < y$, i.e., the limit is a chain on $[y_C, 1]$ placed above an antichain on $[0, y_C]$.

Denote by B_n the partial order induced by P_n on the set $[0, \lfloor y_C n \rfloor]$ and denote by E_n the partial order induced by P_n on the set $[\lfloor y_C n \rfloor + 1, n-1]$. Since a sample of points from P_n can be thought of as a pair of samples, one from B_n and one from E_n, and since we are trying to show that the continuum limit is a chain above an antichain, it is enough to show that

(1) $\lambda(R; B_n) \to \lambda(R; P)$ as $n \to \infty$ for all finite partial orders R;

(2) $\lambda(R; E_n) \to \lambda(R; Q)$ as $n \to \infty$ for all finite partial orders R; and

(3) $\mathbb{E}\#\{i, j \in P_n : i \le y_C n < j \text{ and } i, j \text{ incomparable}\}/n^2 \to 0$ as $n \to \infty$;

where P is the antichain on $[0, y_C]$ and Q is the chain on $[y_C, 1]$. In fact, since Q is a chain, so that $\lambda(R; Q) = 0$ unless R is a chain, by Proposition 2.3 it is enough to show that $\lambda(A_2; E_n) \to 0$ as $n \to \infty$ and so we can show both (2) and (3) by showing that $\mathbb{E}\#\{i < j \in P_n : j > y_C n \text{ and } i, j \text{ incomparable}\}/n^2 \to 0$ as $n \to \infty$. Likewise, since P is an antichain, we only need to show that $\lambda(C_2; B_n) \to 0$ as $n \to \infty$. That is, it is enough to show

(a) $\mathbb{E}\#\{i < j \in P_n : j > y_C n \text{ and } i, j \text{ incomparable}\}/n^2 \to 0$ as $n \to \infty$, and

(b) $\mathbb{E}\#\{i < j \in P_n : j \le y_C n \text{ and } i, j \text{ comparable}\}/n^2 \to 0$ as $n \to \infty$.

Fix $\varepsilon > 0$ with $\varepsilon < y_C/3$. To prove (a), we show that for sufficiently large n, the expected number of incomparable pairs of elements $i < j \in P_n$ with $j > y_C n$ is less than $5\varepsilon n^2$. By Lemma 6.1 we have

$$r_n(z) \ge y_C r_n(y_C)/(1+\varepsilon)z \ge (1-\varepsilon) y_C r_n(y_C)$$

for all $z \in (y_C, 1]$ for sufficiently large n. Since $r_n(y_C) \to \infty$ as $n \to \infty$ we have that $r_n(y_C) > 2/\varepsilon y_C$ for sufficiently large n. Therefore, $r_n(z) > 2(1-\varepsilon)/\varepsilon$ for sufficiently large n, for all $z > y_C$.

So, for $j \ge (y_C + \varepsilon)n$ and $j - i > \varepsilon n$ we have

$$d^{(n)}(i,j) \ge d^{(n)}(j - \lceil \varepsilon n \rceil, j) = \sum_{w=j-\lceil \varepsilon n \rceil + 1}^{j} \rho_w^{(n)} = \frac{1}{n} \sum_{w=j-\lceil \varepsilon n \rceil + 1}^{j} r_n(w/n), \tag{12}$$

and since $j - \lceil \varepsilon n \rceil + 1 > y_C n$, so that $w/n > y_C$ for all $w = j - \lceil \varepsilon n \rceil + 1, \ldots, j$, inequality (12) becomes

$$d^{(n)}(i,j) \geq \frac{1}{n} \sum_{w=j-\lceil \varepsilon n \rceil+1}^{j} 2(1-\varepsilon)/\varepsilon > 2(1-\varepsilon) > 1+\varepsilon.$$

Now, suppose that n is sufficiently large for both the above to hold and so that we can apply Theorems 5.5 and 5.6. We will apply either Theorem 5.5 or 5.6 depending on the size of $\mathbb{E}|D_j|$, for each $j \geq (y_C + \varepsilon)n$. If $\mathbb{E}|D_j| \leq (\log n)^{5/4}$ then we apply Theorem 5.5 and the expected number of i with $\varepsilon n < i < j - \varepsilon n$ and i incomparable to j is at most εn. If $\mathbb{E}|D_j| > (\log n)^{5/4}$ then by definition $j > z_C$. If also $j > z_C + \varepsilon n$ then Theorem 5.6 implies that the expected number of i with $\varepsilon n < i < j - \varepsilon n$ and i incomparable to j is at most εn. Combining this information we have that for all $j > y_C n$, except at most $2\varepsilon n$ (those j in $[y_C n, (y_C + \varepsilon)n] \cup [z_C, z_C + \varepsilon n]$), the expected number of elements i earlier than and incomparable to j is at most $3\varepsilon n$, and therefore the expected number of incomparable pairs $i, j \in P_n$ with $j > y_C n$ is less than $3\varepsilon n \times n + n \times 2\varepsilon n = 5\varepsilon n^2$.

We now prove (b). We have that $\displaystyle\int_{\varepsilon}^{y_C-\varepsilon} R(t)dt \leq 1$. We first show that integrating over the smaller range $[\varepsilon, y_C - 2\varepsilon]$ gives a value strictly less than one, i.e., there exists $\delta < \varepsilon$ such that $\displaystyle\int_{\varepsilon}^{y_C-2\varepsilon} R(t)dt < 1 - 2\delta$.

Consider $I = \displaystyle\int_{y_C-2\varepsilon}^{y_C-\varepsilon} R(t)dt$. If $I = 0$, then since $R(t) \geq 0$ on $[0, y_C)$, we have $R(y_C - 2\varepsilon) = 0$. But then Lemma 6.1 implies that $R(t) = 0$ for all $t \in [\varepsilon, y_C - 2\varepsilon]$ and so $\int_{\varepsilon}^{y_C-2\varepsilon} R(t)dt = 0$ and is certainly less than $1 - 2\delta$ for some $\delta < \varepsilon$.

Otherwise, $I > 0$, so there exists $\delta < \varepsilon$ with $I > 2\delta$ and therefore $\int_{\varepsilon}^{y_C-2\varepsilon} R(t)dt < 1 - 2\delta$ as required.

Now, from Lemma 6.3 we have

$$\limsup_{n\to\infty} \int_{\varepsilon}^{y_C-2\varepsilon} r_n(t)dt \leq \int_{\varepsilon}^{y_C-2\varepsilon} R(t)dt < 1 - 2\delta.$$

Then, by Lemma 6.2,

$$\limsup_{n\to\infty} d^{(n)}(\lceil \varepsilon n \rceil, \lceil (y_C - 2\varepsilon)n \rceil) \leq \limsup_{n\to\infty} \int_{\varepsilon}^{y_C-2\varepsilon} r_n(t)dt$$

and therefore, $d^{(n)}(\varepsilon n, (y_C - k\varepsilon)n) < 1 - \delta$ for all sufficiently large n. Now we apply Theorem 5.5: for any $i, j \in [\varepsilon n, (y_c - 2\varepsilon)n]$, we certainly have $r_n(j/n)$ bounded, and therefore $\mathbb{E}|D_j| = O(\log n)$, so the conditions of Theorem 5.5 hold and $\mathbb{P}(i < j \text{ in } P(\mathbf{t})) < \delta < \varepsilon$. Therefore the expected number of comparable elements $i < j$ with $j \leq y_C n$ is at most $2\varepsilon n \times n + n \times 3\varepsilon n = 5\varepsilon n^2$.

Now we prove case (i). Define $y_C = \inf\{y : \lim_{n\to\infty} r_n(y) = \infty\}$ and set $y_C = 1$ if no such y exists. As in case (ii), we have $\lim_{n\to\infty} r_n(x) = \infty$ for all $x \in (y_C, 1]$. Therefore, the order $\prec$ on the continuum limit T_r has $x \prec y$ for all $y > y_C$ and $x < y$, as in case (ii). Again, it is enough to show

(1) $\lambda(R; B_n) \to \lambda(R; P)$ as $n \to \infty$ for all finite partial orders R;

(2) $\lambda(R; E_n) \to \lambda(R; Q)$ as $n \to \infty$ for all finite partial orders R; and

(3) $\mathbb{E}\#\{i,j \in P_n : i \le y_C n < j \text{ and } i,j \text{ incomparable}\}/n^2 \to 0$ as $n \to \infty$;

where here P is the partially ordered measure space T_r restricted to $[0, y_C]$ and Q is the chain on $[y_C, 1]$. So, again (2) and (3) can be proved by showing that $\mathbb{E}\#\{i < j \in P_n : j > y_C n \text{ and } i,j \text{ incomparable}\}/n^2 \to 0$ as $n \to \infty$, and this follows from exactly the same proof as in case (ii).

It remains to show that $\lambda(R; B_n) \to \lambda(R; P)$ as $n \to \infty$ for all finite partial orders R, for $P = ([0, y_C], \mathcal{B}, \mu_L, \prec)$, where $\prec$ is defined by $x \prec y$ if and only if $\int_x^y r(t)dt > 1$. The proof of this roughly corresponds to the non-trivial case (ii) of Theorem 4.2.

Fix $\varepsilon > 0$, with $\varepsilon < 1/8$, and assume throughout that n is sufficiently large. For each n, take a random order P_n according to $\mathcal{P}_n$, and consider B_n the order induced by P_n on $[0, \lfloor y_C n \rfloor]$. Define an order $\prec_n$ on $[0, y_C]$, by dividing $[0, y_C]$ into $\lfloor y_C n \rfloor$ intervals of length $1/n$, identifying $[i/n, (i+1)/n)$ with $i \in [0, \lfloor y_C n \rfloor]$, and putting $[i/n, (i+1)/n)$ below $[j/n, (j+1)/n)$ if and only if $i < j$ in B_n. Let $\bar{B}_n$ be the atomless partially ordered measure space $([0, y_C], \mathcal{B}, \mu_L, \prec_n)$. A similar argument to that in the proof of Theorem 4.2 implies that $\mathbb{E}\lambda(Q; B_n) - \mathbb{E}\lambda(Q; \bar{B}_n)$ tends to zero as n tends to infinity, for all finite partial orders Q. So, it is enough to show that $\mathbb{E}\lambda(Q; \bar{B}_n) \to \lambda(Q; P)$, which follows if $\mathbb{P}(\prec_n$ induces different partial order from $\prec) \to 0$ as $n \to \infty$. As in the proof of Theorem 4.2, it is enough to consider to elements x, y chosen uniformly at random from $[0, y_C]$ and show that

$$\mathbb{P}(\prec_n \text{ induces different partial order from } \prec \text{ on } \{x,y\}) \to 0 \tag{13}$$

as $n \to \infty$.

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

(i) $\displaystyle\int_{(i+1)/n}^{j/n} r(t)dt > 1$ and i, j are comparable in P_n, or

(ii) $\displaystyle\int_{i/n}^{(j+1)/n} r(t)dt < 1$ and i, j are incomparable in P_n,

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 (13), 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 $1 - \varepsilon \le d^{(n)}(i,j) \le 1 + \varepsilon$ are bad. Also, we assume that all pairs $i < j$ with either $i \notin [\varepsilon n, (y_C - \varepsilon)n]$ or $j \notin [\varepsilon n, (y_C - \varepsilon)n]$ are bad. We need to show that the number of such pairs is small, but first we show that almost all the remaining pairs, i.e., the pairs $i < j$ with $i, j \in [\varepsilon n, (y_C - \varepsilon)n]$ and $|d^{(n)}(i,j) - 1| > \varepsilon$, are good, as follows.

Since n_0 is sufficiently large, Lemmas 6.2 and 6.3 imply that

$$\left| d^{(n)}(\lceil xn \rceil, \lceil yn \rceil) - \int_x^y r(t)dt \right| < \varepsilon/2 \tag{14}$$

for all $n \ge n_0$. Consider a pair $i < j$ with $i, j \in [\varepsilon n, (y_c - \varepsilon)n]$ and $d^{(n)}(i,j) > 1 + \varepsilon$. Equation (14) implies that $\displaystyle\int_{i/n}^{j/n} r(t)dt > 1 + \varepsilon/2$ and since $r(t)$ is finite almost everywhere on $[\varepsilon, y_C - \varepsilon]$, we have $\displaystyle\int_{(i+1)/n}^{j/n} r(t)dt > 1$ for large enough n. So, such a pair is bad if i, j are incomparable in P_n. However, since $r(j/n)$ is almost surely finite, we have $\mathbb{E}|D_j| = O(\log n)$ and so Theorem 5.5 implies that i, j are incomparable in P_n with probability less than ε. Similarly, pairs $i < j$ with $i, j \in [\varepsilon n, (y_c - \varepsilon)n]$ and $d^{(n)}(i, j) < 1 - \varepsilon$, have $\displaystyle\int_{i/n}^{(j+1)/n} r(t)dt < 1$, and i, j are comparable in P_n with probability less than ε. So in total, the expected number of bad pairs $i < j$ with $i, j \in [\varepsilon n, (y_c - \varepsilon)n]$ and $|d^{(n)}(i, j) - 1| > \varepsilon$ is at most εn^2.

Clearly, the number of pairs $i < j$ with either $i \notin [\varepsilon n, (y_C - \varepsilon)n]$ or $j \notin [\varepsilon n, (y_C - \varepsilon)n]$ is at most $2\varepsilon n^2$, and so it remains to show that the number of pairs $i < j \in [\varepsilon n, (y_C - \varepsilon)n]$ with $|d^{(n)}(i, j) - 1| \leq \varepsilon$ is a small fraction of n^2.

Note that by equation (14), it is enough to show that the number of pairs $i < j \in [\varepsilon n, (y_C - \varepsilon)n]$ with

$$\left| \int_{i/n}^{j/n} r(t)dt - 1 \right| \leq 2\varepsilon \tag{15}$$

in small.

Define $G \subseteq [\varepsilon, y_C - \varepsilon]$ as the set of x such that $\int_x^{y_C - \varepsilon} r(t)dt \geq 1 - 2\varepsilon$. Clearly G is an interval of $[\varepsilon, y_C - \varepsilon]$; indeed, it is a down-set (under the natural ordering of $\mathbb{R}$). We only need to consider pairs $i < j$ with $i/n \in G$: if $i/n \notin G$ then

$$\int_{i/n}^{j/n} r(t)dt \leq \int_{i/n}^{y_C - \varepsilon} r(t)dt < 1 - 2\varepsilon,$$

for all $j \in [i + 1, (y_C - \varepsilon)n]$. So we may assume that G is non-empty, and therefore that $\varepsilon \in G$.

For $x \in G$ define the function $g(x)$ implicitly as

$$\int_x^{g(x)} r(t)dt = 1 - 2\varepsilon.$$

Clearly g is increasing on G, and maps G surjectively onto $[g(\varepsilon), y_C - \varepsilon]$.

Now, take i with $i/n \in G$, so that $\int_{i/n}^{g(i/n)} r(t)dt = 1 - 2\varepsilon$. By Lemma 6.1, $t\,r(t) < (1 + \varepsilon)g(\frac{i}{n})r\big(g(\frac{i}{n})\big)$ for $t \in [\frac{i}{n}, g(\frac{i}{n})]$, so

$$1 - 2\varepsilon = \int_{i/n}^{g(i/n)} r(t)dt \leq (1 + \varepsilon)g(\tfrac{i}{n})r\big(g(\tfrac{i}{n})\big) \int_{i/n}^{g(i/n)} \frac{dt}{t} \leq (1 + \varepsilon)g(\tfrac{i}{n})r\big(g(\tfrac{i}{n})\big)\frac{n}{i},$$

or $g(\frac{i}{n})r\big(g(\frac{i}{n})\big) > 2i/3n$, since $\varepsilon < 1/8$. Then, again by Lemma 6.1,

$$\begin{aligned} \int_{g(i/n)}^{g(i/n)+7\varepsilon n/i} r(t)dt &\geq (1 - \varepsilon)g(\tfrac{i}{n})r\big(g(\tfrac{i}{n})\big) \int_{g(i/n)}^{g(i/n)+7\varepsilon n/i} \frac{dt}{t} \\ &\geq (1 - \varepsilon)g(\tfrac{i}{n})r\big(g(\tfrac{i}{n})\big)\frac{7\varepsilon n}{i} > (1 - \varepsilon)\frac{2i}{3n}\frac{7\varepsilon n}{i} > 4\varepsilon. \end{aligned}$$

Therefore,

$$\int_{i/n}^{g(i/n)+7\varepsilon n/i} r(t)dt > 1 + 2\varepsilon,$$

so, for this i, in order for j to satisfy (15), we must have $g(i/n) \leq j/n < g(i/n) + 7\varepsilon n/i$. That is, there are at most $7\varepsilon n^2/i$ points satisfying (15).

So, the total number of pairs is at most

$$\sum_{i\in G}\frac{7\varepsilon n^2}{i}\le 7\varepsilon n^2\sum_{i=\varepsilon n}^{(y_C-\varepsilon)n}\frac{1}{i}\le 7\varepsilon n^2\int_{\varepsilon n-1}^{(y_C-\varepsilon)n}\frac{dx}{x}\le 7\varepsilon n^2\log(1/\varepsilon).$$

Since $\varepsilon\log(1/\varepsilon)\to 0$ as $\varepsilon\to 0$, we have an arbitrarily small fraction of n^2 pairs satisfying (15), as required, and this completes the proof. □

Note that despite the extra freedom of having a sequence of parameters $\mathbf{t}^{(n)}$ for each n, rather than the single parameter $p(n)$ in the random graph order case, the structure of the continuum limits exhibited in Theorem 6.4 are qualitatively very similar to those of an random graph order in that they are all semiorders. Unfortunately, we do not have an analogue to Theorem 4.5 and so we cannot say whether this is the only way for a sequence of classical sequential growth models to have a continuum limit. Nevertheless, we have shown that any other continuum limit has essentially the same structure, in that it is an almost-semiorder.

We finish with an example of a sequence of classical sequential growth models other than a sequence of random graph orders that has a non-trivial continuum limit. The example is not particularly exciting, but given its simplicity we feel that any number of increasingly complex examples could be produced in a similar way. Basically, we define the sequences $\mathbf{t}^{(n)}$ to have only two non-zero terms (excluding the non-zero t_0)—meaning that almost surely the size $|D_{xn}|$ is one of two values—in such a way that the limit of the expectation $\lim_{n\to\infty}\mathbb{E}|D_{xn}|$ exhibits a sharp phase transition between these two values as x varies over $[0,1]$.

For any $0<\alpha<\beta$, and $0<\gamma<1$, define the function $r_{\alpha,\beta,\gamma}$ by

$$r_{\alpha,\beta,\gamma}(x)=\begin{cases}\alpha/x & \text{for } 0<x<\gamma,\\ \beta/x & \text{for } \gamma<x\le 1.\end{cases}$$

Proposition 6.5. *The partially ordered measure space $T_{r_{\alpha,\beta,\gamma}}$ is the continuum limit of the sequence $(\mathcal{P}(\mathbf{t}^{(n)}))_{n=1}^{\infty}$ of classical sequential growth models, where the only non-zero terms of $\mathbf{t}^{(n)}$ are*

$$t_0^{(n)}=1,\quad t_{\alpha\log n}^{(n)}=\binom{\gamma n}{\beta\log n},\quad t_{\beta\log n}^{(n)}=\binom{\gamma n}{\alpha\log n},$$

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

Proof. As alluded to earlier, we show that $\mathbb{E}|D_{\gamma n+l}|$ is asymptotically equal to $\alpha\log n$ when $l\le -n/\sqrt{\log n}$, and asymptotically equal to $\beta\log n$ when $l\ge n/\sqrt{\log n}$.

Recall that in general

$$\mathbb{E}|D_y|=\frac{\sum_{k=0}^{y}kt_k^{(n)}\binom{y}{k}}{\sum_{k=0}^{y}t_k^{(n)}\binom{y}{k}},$$

so that here

$$\begin{aligned}\mathbb{E}|D_{\gamma n+l}|&=\frac{\alpha\log n\binom{\gamma n}{\beta\log n}\binom{\gamma n+l}{\alpha\log n}+\beta\log n\binom{\gamma n}{\alpha\log n}\binom{\gamma n+l}{\beta\log n}}{1+\binom{\gamma n}{\beta\log n}\binom{\gamma n+l}{\alpha\log n}+\binom{\gamma n}{\alpha\log n}\binom{\gamma n+l}{\beta\log n}}\\ &\sim\frac{\alpha\log n+R\beta\log n}{1+R},\end{aligned}$$

where

$$R=\frac{\binom{\gamma n}{\alpha\log n}\binom{\gamma n+l}{\beta\log n}}{\binom{\gamma n}{\beta\log n}\binom{\gamma n+l}{\alpha\log n}}\sim\left(1+\frac{l}{\gamma n}\right)^{(\beta-\alpha)\log n}\sim\exp\left(\frac{l(\beta-\alpha)\log n}{\gamma n}\right).$$

Since $R \to 0$ if $l \le -n/\sqrt{\log n}$, and $R \to \infty$ if $l \ge n/\sqrt{\log n}$, we have,

$$\mathbb{E}|D_{\gamma n+l}| \sim \begin{cases} \alpha \log n & \text{if } l \le -n/\sqrt{\log n}, \\ \beta \log n & \text{if } l \ge n/\sqrt{\log n}. \end{cases}$$

Therefore the limit $\lim_{n\to\infty} r_n(x)$ exists for all $x \in (0,1] \setminus \{\gamma\}$ and is equal to $r_{\alpha,\beta,\gamma}$. So, by Theorem 1.8, the continuum limit exists, and it is equal to $T_{r_{\alpha,\beta,\gamma}}$. □

References

[1] M.H.Albert and A.M.Frieze, Random graph orders, Order **6** (1989), no. 1, 19–30.

[2] N.Alon, B.Bollobás, G.Brightwell and S.Janson, Linear extensions of a random partial order, Ann. Appl. Probab. **4** (1994), 108–123

[3] N.Alon, J.H.Spencer and P.Erdős, *The Probabilistic Method*, John Wiley & Sons, 1992.

[4] B.Bollobás and G.Brightwell, Box-spaces and random partial orders, Trans. Amer. Math. Soc. **324** (1991), no. 1, 59–72.

[5] B.Bollobás and G.Brightwell, The width of random graph orders, Math. Sci. **20** (1995), no. 2, 69–90.

[6] 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.

[7] B.Bollobás and G.Brightwell, The structure of random graph orders, SIAM J. Discrete Math. **10** (1997), no. 2, 318–335.

[8] P.C.Fishburn, *Interval Orders And Interval Graphs*, John Wiley & Sons, 1985.

[9] N.Georgiou, The random binary growth model, Random Structures Algorithms **27** (2005), no. 4, 520–552.

[10] G.H.Hardy, J.E.Littlewood and G.Pólya, *Inequalities*, Cambridge University Press, 1934, 1952.

[11] J.H.Kim and B.Pittel, On tail distribution of interpost distance, J. Combin. Theory Ser. B **80** (2000), no. 1, 49–56.

[12] L.Lovász and B.Szegedy, Limits of dense graph sequences, J. Combin. Theory Ser. B **96** (2006), no. 6, 933–957.

[13] B.Pittel and R.Tungol, A phase transition phenomenon in a random directed acyclic graph, Random Structures Algorithms **18** (2001), no. 2, 164–184.

[14] D.P.Rideout and R.D.Sorkin, Classical sequential growth dynamics for causal sets, Phys. Rev. D (3) **61** (2000), no. 2, 024002, 16 pp.

[15] 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, 15 pp.

[16] K.Simon, D.Crippa and F.Collenberg, On the distribution of the transitive closure in a random acyclic digraph, Algorithms—ESA '93 (Bad Honnef, 1993), 345–356, Lecture Notes in Comput. Sci., 726, Springer, Berlin, 1993.