The Length of Random Subsets of Boolean Lattices*

to appear in *Random Structures and Algorithms*

Y. Kohayakawa,†
Instituto de Matemática e Estatística, Universidade de São Paulo
Rua do Matão 1010, 05508–900 São Paulo, Brazil
E-mail address: `yoshi@ime.usp.br`

B. Kreuter, and D. Osthus‡
Institut für Informatik, Humboldt-Universität zu Berlin
Unter den Linden 6, 10099 Berlin, Germany
E-mail address: `{kreuter,osthus}@informatik.hu-berlin.de`

Abstract

We form the random poset $\mathcal{P}(n,p)$ by including each subset of $[n] = \{1,\dots,n\}$ with probability p and ordering the subsets by inclusion. We investigate the length of the longest chain contained in $\mathcal{P}(n,p)$. For $p \geq e/n$ we obtain the limit distribution of this random variable. For smaller p we give thresholds for the existence of chains which imply that almost surely the length of $\mathcal{P}(n,p)$ is asymptotically $n(\log n - \log\log 1/p)/\log 1/p$.

1 Introduction and basic definitions

We investigate random subsets of the power set $\mathcal{P}(n)$ of $[n] = \{1,\dots,n\}$. Such a random subset, denoted $\mathcal{P}(n,p)$, is obtained by selecting each element of $\mathcal{P}(n)$ with probability p independently of all other elements. If we order the elements of $\mathcal{P}(n)$ by inclusion, we obtain a Boolean lattice. If we order the elements of $\mathcal{P}(n,p)$ by inclusion, we obtain a partially ordered set (or simply *poset*) which is a subposet of the above Boolean lattice.

Just as the theory of random graph tells us what a "typical" graph looks like, we can say what the properties of a "typical" subset of $\mathcal{P}(n)$ are if we know what $\mathcal{P}(n,p)$ looks like. Basic parameters of a graph are for instance the clique number and the chromatic number, whereas for $\mathcal{P}(n,p)$ the most fundamental one is certainly its length.

Here the *length* of $\mathcal{P}(n,p)$, for which we write length $\mathcal{P}(n,p)$, is defined to be the length of the maximum chain contained in $\mathcal{P}(n,p)$. A *chain* is a subset

*Research supported by PROBRAL project number 026/95, a CAPES–DAAD exchange programme.

†Partially supported by FAPESP (Proc. 96/04505–2), by CNPq (Proc. 300334/93–1), and by MCT/ FINEP (PRONEX project 107/97).

‡Graduate school "Algorithmische Diskrete Mathematik", supported by Deutsche Forschungsgemeinschaft grant GRK/3-98

of $\mathcal{P}(n)$ whose elements are pairwise comparable and the *length* of a chain is its cardinality minus one. Rényi was the first to investigate the behaviour of this random variable: answering a question of Erdős, in [12] he determined the threshold for the existence of a chain of length one in $\mathcal{P}(n,p)$. Recently, Kreuter [10] obtained (amongst other results) thresholds for the appearance of chains of any fixed length.

This raises the question how the length evolves if we increase p beyond the range considered in [10]. In this paper we answer this question by studying the behaviour of certain "regular" chains in $\mathcal{P}(n,p)$.

Another important parameter of $\mathcal{P}(n,p)$ is the cardinality of its maximum antichain, also called the *width*. This is studied in [9, 11], where the results proven in [11] are based on the central lemma of this paper, Lemma 7. Indeed, somewhat surprisingly it turns out that Lemma 7 is strong enough to bound the width of $\mathcal{P}(n,p)$ by double counting chains in $\mathcal{P}(n,p)$, following the approach of Lubell (see e.g. [1]) in his elegant proof of Sperner's lemma. For large p, this seems to be the only way to obtain these results.

Brightwell [6] gives an overview of some other models of random posets. For example, several authors (see e.g. Bollobás and Brightwell [3]) have studied the length of the random poset generated by picking n random points in the k-dimensional unit cube. However this model, although superficially similar to ours, produces different results and requires other proof techniques. Our methods are more related to those used in the past to investigate the diameter and other parameters of random subgraphs of the n-dimensional hypercube (see e.g. Bollobás, Kohayakawa and Łuczak [4, 5]). Also, Fill and Pemantle [7] prove the analogue of Theorem 4 in the context of oriented bond percolation on the hypercube.

Throughout, we refer to the subset $\mathcal{P}_t(n) \subset \mathcal{P}(n)$ whose elements have cardinality t as the tth *level* of $\mathcal{P}(n)$. The levels of $\mathcal{P}(n,p)$ are defined in the same way. The (Hamming) distance $d_H(x,y)$ of x, $y \in \mathcal{P}(n)$ is given by $|(x\setminus y)\cup(y\setminus x)|$. We write *a.s.* to mean "with probability tending to one as n tends to infinity".

This paper is organized as follows. In the next section, we state our main results, Theorems 1 and 4. Our main technical tool is given in Section 3, where we state and prove Lemma 7. Theorems 1 and 4 are then proved in Sections 4 and 5.

2 Results

Kreuter [10] proved that the length of $\mathcal{P}(n,p)$ is constant when p decreases exponentially. The following theorem tells us about the length of $\mathcal{P}(n,p)$ when p decays at a rate slower than exponential:

Theorem 1 *Let $k = k(n) \leq n-1$ be an integer valued function tending to infinity. Then*

$$\lim_{n\to\infty} \mathbb{P}(\,\mathrm{length}\,\mathcal{P}(n,p) \geq k-1\,) = \begin{cases} 0 & \text{if } p\,(k+1)^{n/k} \to 0, \\ 1 & \text{if } p\,(k+1)^{n/k}(k/n)^2 \to \infty. \end{cases} \tag{1}$$

At the end of Section 3 we indicate how one may remove the factor $(k/n)^2$ at the expense of more extensive calculations. One consequence of Theorem 1 is that when viewed as a poset, $\mathcal{P}(n,p)$ is in a sense close to being a ranked poset. More precisely, given a constant ν with $0<\nu<1$ and an element $x \in \mathcal{P}_{\nu n}(n,p)$ one can apply Theorem 1 to $\mathcal{P}(\nu n)$ to obtain bounds on the length of the longest chain in $\mathcal{P}(n,p)$ whose maximal element is x.

An immediate corollary of the above theorem is that when p decreases polynomially in n, the degree of the polynomial turns out to be the average distance between two elements of a typical maximum chain.

Corollary 2 *Let $r \geq 1$ be a fixed constant. Then*

$$\lim_{n\to\infty} \mathbb{P}\Big(\operatorname{length}\mathcal{P}(n,p) \geq \frac{n}{r} - 2\Big) = \begin{cases} 0 & \textit{if } pn^r \to 0 \\ 1 & \textit{if } pn^r \to \infty. \end{cases} \tag{2}$$

We remark that the above threshold could be sharpened slightly by the use of branching processes and indicate how this may be done at the end of the final section. Another immediate corollary of Theorem 1 is the following asymptotic formula for the length of $\mathcal{P}(n,p)$.

Corollary 3 *If $c^n = o(p)$ for any $0<c<1$ and $p = \mathcal{O}(1/n)$, then* a.s.

$$\operatorname{length}\mathcal{P}(n,p) = (1+o(1))\,\frac{n}{\log 1/p}\,(\log n - \log\log 1/p)\,. \tag{3}$$

Our second theorem (which improves Corollary 2 for the special case when $r=1$) gives a detailed picture of the final stage of the evolution of the length of $\mathcal{P}(n,p)$:

Theorem 4 *Let $c \geq e$ be a given constant. Define $p(n) = c/n$. Let $\eta = \eta(c)$ satisfy $\eta = e^{c(\eta-1)}$ and $0<\eta<1$. Then*

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}(\operatorname{length}\mathcal{P}(n,p) = n-2) = (1-\eta)^2, \\
&\lim_{n\to\infty} \mathbb{P}(\operatorname{length}\mathcal{P}(n,p) = n-3) = 2\eta(1-\eta), \\
&\lim_{n\to\infty} \mathbb{P}(\operatorname{length}\mathcal{P}(n,p) = n-4) = \eta^2.
\end{aligned}$$

We remark that η is the extinction probability of a branching process whose family size has Poisson distribution with parameter c. Also note that since the top and bottom levels of $\mathcal{P}(n)$ contain only a single element, *a.s.* we have that $\operatorname{length}\mathcal{P}(n,p) < n-1$ as long as $p \to 0$. Finally, the following proposition shows that the previous theorem is essentially best possible in the sense that if $c<e$, then $\mathcal{P}(n,p)$ becomes significantly shorter. Indeed, if $p = p(n) = c/n$, Theorem 1 and Proposition 5 imply that, very crudely, we have

$$n - \operatorname{length}\mathcal{P}(n,p) \quad \begin{cases} = \mathcal{O}(1) & \text{if } c \geq e \\ \to \infty & \text{if } c < e \end{cases} \tag{4}$$

almost surely. We believe that the fact that the techniques in this paper are enough to identify the sharp change of behaviour given in (4) is quite pleasant.

Proposition 5 *Let $c < e$ be a given constant. Define $p(n) = c/n$. Then* a.s.

$$\operatorname{length} \mathcal{P}(n,p) \leq \left(1 - \frac{\log(e/c)}{3\log n}\right) n.$$

Proof. Let

$$a = \left\lceil \frac{\log(e/c)}{3\log n} n \right\rceil.$$

For $\mathcal{Q} \subset \mathcal{P}(n)$ denote by $Y(\mathcal{Q})$ the number of $(n-a)$-element chains in $\mathcal{Q}$. A chain on $n-a$ elements corresponds to an ordered partition of $[n]$ into $n-a$ nonempty sets. To bound the number of those, first choose $n-a$ elements, one for each partition class. There are $(n)_{n-a} = n!/a!$ ways of doing this. Then order the remaining a elements and distribute them among the $n-a$ classes, with $(n-a)^a$ possibilities for this second step. Hence by Stirling's formula

$$Y(\mathcal{P}(n)) \leq n^a n! = \mathcal{O}(n^{a+1/2}(n/e)^n).$$

Thus the expected value of $Y = Y(\mathcal{P}(n,p))$ is

$$\mathbb{E}(Y) = \mathcal{O}(n^{a+1/2}(n/e)^n p^{n-a}) = \mathcal{O}(n^{2a+1/2}(c/e)^{n-a}) = o(1)$$

by the definition of a. □

3 Regular chains

The aim of this section is to prove Lemma 7 below. This lemma implies that the number of certain "regular" chains is not too far from its expected value provided that p is large enough. Once we have this result, Theorem 1 will follow immediately, while we still have to do some additional work for Theorem 4. These "regular" chains are defined as follows: suppose x, $y \in \mathcal{P}(n)$ are given with $x \subset y$. Given any r and k, we say $(x_1, \ldots, x_k)$ is an (x,y,r,k)-*chain* if $x \subset x_1 \subset \cdots \subset x_k \subset y$ with the additional condition that $|x_i| = |x| + \lceil ir \rceil$ for all i. Since r and k are always related in what follows, (x,y,r,k)-chains are simply referred to as (x,y,r)-chains.

Throughout this section we assume that $x \in \mathcal{P}_t(n)$ and $y \in \mathcal{P}_{n-t}(n)$ for some fixed integer t. Also k, the number of elements in a chain, is a given integer function of n with $k \leq n-2t-1$. As a measure of the average distance between two chain elements we define

$$r = \begin{cases} 1 & \text{if } k = n-2t-1 \\ (n-2t-1)/(k+1) & \text{if } k < n-2t-1. \end{cases} \tag{5}$$

The extra minus one in the fraction is a technical artifice to ensure that $\mathcal{P}_{|x_i|}(n,p)$ is large enough even when r is not an integer and i is close to k. We separate the two cases since we want to achieve the best possible result when $r = 1$. Now, for given $m \in \mathbb{N}$, let

$$\psi(m) = \begin{cases} m+1 & \text{if } k = n-2t-1 \\ \lfloor (m+1)r \rfloor + 1 & \text{if } k < n-2t-1. \end{cases} \tag{6}$$

Writing $x_0 = x$ and $x_{k+1} = y$, the Hamming distance between chain elements x_j and $x_{j'}$ with $j' > j$ is then at most $\psi(j' - j - 1)$. We shall often make use of the fact that $\psi(j') - \psi(j) \leq (j' - j)r + 1$. Finally we assume that

$$p(n) = d\,\psi(0)!\left(\frac{e}{(n-2t)(k+1)^{1/k}}\right)^r \tag{7}$$

with $d = d(n) \geq 1$. The following form of Stirling's formula (valid for all $n > 0$) will suffice for our purposes:

$$\sqrt{n}\left(\frac{n}{e}\right)^n < n! < 3\sqrt{n}\left(\frac{n}{e}\right)^n. \tag{8}$$

Let $\mathcal{S}$ be a family of subsets of $\mathcal{P}(n)$. In what follows, $\mathcal{S}$ will usually consist of various chains in $\mathcal{P}(n)$. For $\alpha \in \mathcal{S}$ let I_α be the indicator of α, i.e., we have $I_\alpha = 1$ iff $\alpha \subseteq \mathcal{P}(n,p)$. Now define

$$X_{\mathcal{S}} = \sum_{\alpha \in \mathcal{S}} I_\alpha. \tag{9}$$

Proposition 6 *Let $t \in \mathbb{N}$, $x \in \mathcal{P}_t(n)$ and $y \in \mathcal{P}_{n-t}(n)$ with $x \subset y$ be given. Let $\mathcal{S} = \mathcal{S}(x,y,r)$ be the set of all (x,y,r)-chains in $\mathcal{P}(n)$. Then*

$$\lim_{n\to\infty} \mathbb{E}(X_{\mathcal{S}}) = \infty.$$

Proof. Let $R_k = |\mathcal{S}|$ denote the number of (x,y,r)-chains in $\mathcal{P}(n)$. By definition of r and ψ we have

$$R_k = \binom{n-2t}{\lceil r\rceil, \lceil 2r\rceil - \lceil r\rceil, \ldots, n-2t-\lceil kr\rceil} \geq \frac{(n-2t)!}{\psi(0)!^{k+1}}.$$

However, $\mathbb{E}(X_{\mathcal{S}}) = R_k\, p^k$. The result now follows by Stirling's formula (checking the cases $r = 1$ and $r > 1$ separately). □

For $\alpha \neq \beta \in \mathcal{S}$ write $\alpha \sim \beta$ if $\alpha \cap \beta \neq \emptyset$. Then we define

$$\Delta(\mathcal{S}) = \sum_{\alpha\sim\beta} \mathbb{E}(I_\alpha I_\beta), \tag{10}$$

where the sum is over all ordered pairs α, $\beta \in \mathcal{S}$ with $\alpha \sim \beta$.

Lemma 7 *Let $t \in \mathbb{N}$, $x \in \mathcal{P}_t(n)$ and $y \in \mathcal{P}_{n-t}(n)$ with $x \subset y$ be given. Let $\mathcal{S} = \mathcal{S}(x,y,r)$ be the set of all (x,y,r)-chains in $\mathcal{P}(n)$. Then*

$$\Delta(\mathcal{S}) \leq \frac{9}{d}\,\mathbb{E}(X_{\mathcal{S}})^2,$$

provided n is large enough.

There are several results on random subgraphs of the hypercube which rely on a similar lemma, essentially giving bounds on the number of pairs of various kinds of paths in the hypercube. In fact, case (a) of the proof of Lemma 7 is similar to Lemma 2.2 proved by Fill and Pemantle [7], but is sketched for completeness. However, in general their bounds do not apply to our model (even for the simplest case $r = 1$) owing to the differences in the geometry of the paths involved.

Proof of Lemma 7. Consider any ordered pair (α, β) of (x, y, r)-chains with $\alpha \sim \beta$. Let $i = i(\alpha, \beta) = |\alpha \cap \beta|$ be the number of common elements. Note that $0 < i < k$.

A *disjoint interval* γ is a subset of $\alpha \setminus \beta$ with the property that, whenever we have w, $w' \in \gamma$, $v \in \alpha$, and $w \subset v \subset w'$, we must have $v \in \gamma$. We then let l be the maximum cardinality (or *size*) of a disjoint interval and let s be the number of disjoint intervals. Note that $0 < l \le k - i$ and $0 < s \le k - i$.

Now write $\sum_{i,l,s}$ for the sum over all ordered pairs of (x, y, r)-chains α and β with $i(\alpha, \beta) = i$, $l(\alpha, \beta) = l$, and $s(\alpha, \beta) = s$. Let

$$\Delta_{ils} = \sum_{i,l,s} \mathbb{E}(I_\alpha I_\beta)$$

and define Δ_i, Δ_{il}, and Δ_{is} similarly. We can then write $\Delta = \Delta(\mathcal{S})$ as

$$\Delta = \sum_{i=1}^{k-1} \Delta_i, \quad \text{where} \quad \Delta_i = \sum_{l=1}^{k-i} \Delta_{il} = \sum_{s=1}^{k-i} \Delta_{is} \quad \text{and} \quad \Delta_{is} = \sum_{l=1}^{k-i} \Delta_{ils}. \tag{11}$$

Let now i, l and $s > 0$ with $i + l \le k$ and $i + s \le k$ be given. Then an *admissible (i, l, s)-configuration* is a set $A \subset [k]$ with $|A| = i$ such that the longest interval of consecutive elements in $[k]\setminus A$ contains exactly l elements and such that the number of intervals of consecutive elements is exactly s. The set of all admissible (i, l, s)-configurations, denoted $\mathcal{A}_{ils}$, corresponds to the possible intersection patterns of two (x, y, r)-chains (indeed, the configurations give the positions in which two chains may intersect). Let $C_{ils} = |\mathcal{A}_{ils}|$. Let $C_{il} = \sum_s C_{ils}$, $C_{is} = \sum_l C_{ils}$, and $C_i = \sum_l C_{il}$. Note the trivial bounds $C_{ils} \le C_{il} \le C_i \le \binom{k}{i}$.

Now we investigate the number of ways in which a given configuration can be *extended.* Let i, l and s be given and let $\alpha = (x_1 \subset x_2 \subset \cdots \subset x_k)$ be a fixed (x, y, r)-chain and $A \in \mathcal{A}_{ils}$. Let $E(A)$ be the number of possible extensions corresponding to the intersection pattern given by A: i.e., $E(A)$ is the number of β that are (x, y, r)-chains in $\mathcal{P}(n)$ with $\alpha \cap \beta = \{x_i \colon i \in A\}$. Let

$$E_{ils} = \max_{A \in \mathcal{A}_{ils}} E(A).$$

This time, define $E_{il} = \max_s E_{ils}$ and $E_{is} = \max_l E_{ils}$. To estimate E_{ils}, note that each disjoint interval, between x_j and $x_{j'}$ with $j' > j$ say, contributes a factor of

$$\binom{\lceil j'r \rceil - \lceil jr \rceil}{\lceil (j+1)r \rceil - \lceil jr \rceil, \ldots, \lceil j'r \rceil - \lceil (j'-1)r \rceil}.$$

By the log-convexity of factorials, $E(A)$ is then maximized if A has as few disjoint intervals as possible and if as many of these intervals as possible have maximum size. Instead of bounding E_{ils} directly, it turns out to be more convenient to consider E_{ils}/R_k, where R_k is defined to be the number of (x, y, r)-chains in $\mathcal{P}(n)$ as in the proof of Proposition 6. This has the advantage that some of the factorials involving $\lceil r \rceil$ and $\lfloor r \rfloor$ cancel straightaway. The same also applies to E_{il} and E_{is}: for instance, the above argument and the formula for R_k given in the proof of Proposition 6 imply that

$$\frac{E_{il}}{R_k} \leq \frac{\psi(0)!^{i-1}}{(n-2t)!} \psi(l)!\, \psi(k-i-l)!. \tag{12}$$

Indeed, note that we must have at least one disjoint interval of size l and that in the worst case, all the remaining disjoint vertices are contained in a single disjoint interval of size $k-i-l$. We remark that the above inequality also shows that one may interpret E_{il}/R_k as an upper bound on the fraction of all $(x, y, 1)$-chains whose intersection with a given (x, y, r)-chain is an admissible (i, l)-configuration.

For notational convenience let $\delta_{ils} = \Delta_{ils}/\mathbb{E}(X_{\mathcal{S}})^2$ and define the δ_i etc. similarly. Recalling that $\mathbb{E}(X_{\mathcal{S}}) = R_k\, p^k$, we have

$$\delta_{ils} = \frac{\Delta_{ils}}{\mathbb{E}(X_{\mathcal{S}})^2} \leq \frac{R_k\, p^{2k-i}\, C_{ils}\, E_{ils}}{\mathbb{E}(X_{\mathcal{S}})^2} \leq \frac{C_{ils}\, E_{ils}}{R_k\, p^i}. \tag{13}$$

We will bound the δ_{ils} by applying different estimates for different ranges of i, l and s. To fix these ranges we need to define the following constants:

$$\lambda = 1/40 \quad \text{and} \quad \rho = 1/40e. \tag{14}$$

Choose $\kappa > 0$ sufficiently small so that

$$(1-\lambda)(1+\kappa)^2 < 1. \tag{15}$$

Choose $\sigma > 0$ sufficiently small so that

$$(e/\sigma)^{2\sigma}(1-\lambda)^{\rho}(1+\kappa)^2 < 1. \tag{16}$$

Finally choose $l_0 \in \mathbb{N}$ large so that for any $l \geq l_0$,

$$2\left(\frac{4e}{l}\right)^{\sigma} < 1 \quad \text{and} \quad \psi(l)!^{1/(l-1)} \leq \left(\frac{lr(1+\kappa)}{e}\right)^{r}. \tag{17}$$

We assume that n is large enough whenever needed. We also make use of the following inequality, which follows immediately from Stirling's formula (8) and the definition of r and p in (5) and (7):

$$\frac{\psi(0)!^i}{p^i\,(n-2t)!} \leq \frac{(k+1)^{ir/k}}{d\sqrt{kr}} \left(\frac{e}{(k+1)r}\right)^{\psi(k-i)}. \tag{18}$$

(a) $i < \lambda k$.

It turns out that the only significant contribution of Δ_{ils} to Δ comes from the case when i is small and $l = k - i$ (which implies that $s = 1$): in the case (a1), when l is small, we have only a trivial bound on the number of configurations but this is compensated by the fact that, in this case, the number of extensions is severely limited. In the case (a2), when l is large, the situation is reversed. Overall this allows us to show that δ_i is a constant multiple of the worst case contribution $\delta_{i,k-i}$.

(a1) $l \le k - 4i$.

We have $\sum_{l=1}^{k-4i} C_{il} \le C_i \le \binom{k}{i}$ and by (12),

$$\frac{E_{il}}{R_k} \le \frac{\psi(0)!^i}{(n-2t)!}\,\psi(k-4i)!\,\psi(3i)!.$$

Thus, using $r \ge 1$ and $\lambda = 1/40$, we have

$$\begin{aligned}
\sum_{l \le k-4i} C_{il}\frac{E_{il}}{R_k} &\le \frac{\psi(0)!^i}{(n-2t)!}\,\psi(k-i)!\,\frac{\psi(k-4i)!}{\psi(k-i)!}\,\psi(3i)!\binom{k}{i} \\
&\le \frac{\psi(0)!^i}{(n-2t)!}\,\psi(k-i)!\left(\frac{(4ir)^3}{((k-4i)r)^3}\,\frac{ek}{i}\right)^{\lfloor ir \rfloor} \\
&\le \frac{\psi(0)!^i}{(n-2t)!}\,\psi(k-i)!\left(\frac{64e\lambda^2}{(1-4\lambda)^3}\right)^{\lfloor ir \rfloor} \\
&\le \frac{\psi(0)!^i}{(n-2t)!}\,\psi(k-i)!\,\frac{1}{4}.
\end{aligned}$$

(a2) $l > k - 4i$.

As there are $k-l+1$ choices for picking the position of the lowest disjoint interval of length l and at most $\binom{k-l-1}{i-1}$ choices for picking the remaining elements we have

$$C_{il} \le (k-l+1)\binom{k-l-1}{i-1}. \tag{19}$$

Define M_{il} to be the product of the right-hand sides of inequalities (19) and (12). Thus $C_{il}\,E_{il}/R_k \le M_{il}$. Furthermore, since $r \ge 1$,

$$\begin{aligned}
\frac{M_{i,l+1}}{M_{i,l}} &\ge \frac{k-l}{k-l+1}\,\frac{k-i-l}{k-l-1}\left(\frac{(l+1)r}{(k-i-l+2)r}\right)^r \\
&\ge \frac{k-i-l}{k-i-l+2}\,\frac{k-4i}{4i} \ge \frac{1}{3}\,\frac{1-4\lambda}{4\lambda} = 3.
\end{aligned}$$

Since $\sum_{j=0}^{\infty} 3^{-j} = 3/2$, the total contribution of the M_{il} for $l > k-4i$ is at most $3/2$ times that of the maximum term $l = k - i$.

Putting both (a1) and (a2) together gives

$$\delta_i \le \sum_{l=1}^{k-i} \delta_{il} \le 2\,\frac{M_{i,k-i}}{p^i} \le 2\,(i+1)\,\frac{\psi(k-i)!}{(n-2t)!}\,p^{-i}$$

But $n-2t-\psi(k-i)\geq ir$ and $i\leq\lambda k$ imply that

$$\frac{\psi(k-i)!}{(n-2t)!}\leq((1-2\lambda)(n-2t))^{-ir}.$$

Thus

$$\delta_i\leq\, 2(i+1)\,d^{-i}\,\left(\frac{(k+1)^{1/k}}{e(1-2\lambda)}\right)^{ir}\leq\, 2(i+1)\left(\frac{1}{2d}\right)^i,$$

where the second line follows since the fraction in brackets is less than 1/2. Thus, using this time that $d\geq 1$,

$$\sum_{1\leq i\leq\lambda k}\delta_i\leq 2\left(\left(1-\frac{1}{2d}\right)^{-2}-1\right)\leq\frac{8}{d}.$$

(b) $\lambda k\leq i<(1-\rho)k$.

(b1) $l<l_0$.

We have $C_{il}\leq\binom{k}{i}\leq 2^k$. To estimate E_{il}/R_k, note that chain elements in the intersection which lie at either end of a disjoint stretch have distance at most $\psi(l)$. Thus for each disjoint interval (of size $\bar{l}$ say) we have a contribution of a factor of $\psi(\bar{l})!\leq\psi(0)!\,(2\psi(l))^{\bar{l}r}$. Using the fact that there are $k-i$ disjoint vertices and at most $i+1$ disjoint intervals yields the following bound:

$$\frac{E_{il}}{R_k}\leq\frac{\psi(0)!^i}{(n-2t)!}\,(2\psi(l))^{(k-i+1)r}.$$

Hence, by inequality (18), and noting (crudely) that $(k+1)^{ir/k}\leq(k+1)^r<e^{\rho kr}$, we have

$$\begin{aligned}
\delta_{il}&\leq\frac{\psi(0)!^i}{p^i(n-2t)!}\,(2\psi(l))^{(k-i+1)r}\,2^k\\
&\leq\frac{e^{\rho kr}}{d}\left(\frac{e}{(k+1)r}\right)^{(k-i+1)r}\quad(2(l_0+1)r)^{(k-i+1)r}\,2^k\\
&\leq\frac{1}{d}\left(\frac{2^{1+1/\rho}(l_0+1)e^2}{k}\right)^{\rho kr}\\
&\leq\frac{1}{dk^3}.
\end{aligned}$$

In the last line we used the fact that $k\to\infty$.

(b2) $l>l_0$.

Fix a chain α and any configuration with parameters i, l, and s. Denote the size of the disjoint intervals by l_j, for $1\leq j\leq s$. Then there are at most $\psi(0)^{i-(s-1)}\prod_{j=1}^s\psi(l_j)!$ $(x,y,1)$-chains which meet α exactly at the i chosen points. However, this product is maximized if we have as many disjoint intervals of size l as possible, say q of them. The remaining disjoint vertices will be distributed among one interval of size $\bar{l}$, with $1 \le \bar{l} < l$, and $s-q-1$ unit size intervals. Then noting that $ql + \bar{l} + (s-q-1) = k-i$ and $s-q-1 < s$ gives

$$\frac{E_{ils}}{R_k} \le \frac{\psi(0)!^{i-(s-1)}}{(n-2t)!}\,\psi(l)!^{k-i-s-\bar{l}/l-1}\,\psi(\bar{l})!\,\psi(1)!^{s-1}$$

$$\le \frac{\psi(0)!^{i}}{(n-2t)!}\,\psi(l)!^{k-i-s/l-1}\,(4r)^{sr}.$$

Now use $l \le k-i < (1-\lambda)k$ and the definition of l in (17) to obtain

$$\begin{aligned}\frac{E_{ils}}{R_k} &\le \psi(0)!^{i}\left(\frac{e}{n-2t}\right)^{n-2t}\left(\frac{lr(1+\kappa)}{e}\right)^{(k-i)r}\left(\frac{4er}{(l+1)r}\right)^{sr}\\ &\le \psi(0)!^{i}\left(\frac{e}{n-2t}\right)^{ir}((1-\lambda)(1+\kappa))^{(k-i)r}\left(\frac{4e}{l}\right)^{s}.\end{aligned}$$

By the definition of p in (7), and again crudely estimating $(k+1)^{ir/k} \le (k+1)^r \le (1+\kappa)^{\rho kr}$, we have

$$\delta_{il} \le \frac{1}{d}\sum_{s=1}^{k} C_{ils}\left((1-\lambda)(1+\kappa)^2\right)^{\rho kr}\left(\frac{4e}{l}\right)^{s}.$$

When bounding C_{ils} we distinguish two cases. First suppose that $s \le \sigma k$. Note that any admissible configuration may be defined in the following way: we first fix the lowest point of each of the s disjoint intervals. $\binom{k}{s}$ is a crude bound on the number of choices. This first step fixes the positions of at least $s-1$ common vertices, since unless a disjoint interval is right at the top, it must obviously be below a common vertex. We now fix a configuration uniquely by successively adding a further $i-s$ common vertices on top of the existing common intervals. In other words, we are now drawing $i-s+1$ elements (with replacement) from a set of s distinct elements, giving us

$$\binom{s+(i-s+1)-1}{i-s+1} = \binom{i}{s-1}$$

choices on the second step. Thus

$$C_{ils} \le \binom{k}{s}\binom{i}{s-1} \le \left(\frac{ek}{s}\right)^{s}\left(\frac{ei}{s-1}\right)^{s-1} \le \left(\frac{ek}{s}\right)^{2s} \le \left(\frac{e}{\sigma}\right)^{2\sigma k}.$$

For $s > \sigma k$ we just use the trivial bound $C_{ils} \le 2^k$. Plugging in and using (16) and (17) this yields

$$\begin{aligned}\delta_{il} &\le \frac{1}{d}\left(\sum_{1\le s\le \sigma k}\left(\frac{e}{\sigma}\right)^{2\sigma k}((1-\lambda)(1+\kappa)^2)^{\rho kr} + \sum_{\sigma k < s \le k} 2^k\left(\frac{4e}{l}\right)^{s}\right)\\ &\le \frac{1}{d}\left(k^{-4} + k2^k\left(\frac{4e}{l_0}\right)^{\sigma k}\right)\\ &\le \frac{1}{dk^3}. \qquad (20)\end{aligned}$$

(c) $i \geq (1-\rho)k$.

There are at most $\binom{i+s}{s}$ ways of choosing the relative order of i common vertices and the s disjoint intervals. We now fix the size of the intervals, by distributing the $k-i-s$ remaining "disjoint" vertices among the s disjoint intervals. We are choosing these elements (with replacement) from s "urns", giving at most $\binom{k-i-1}{k-i-s}$ possibilities for this step. Hence, by applying Stirling's inequality and then that $i+s \leq k$ and $m^{-m} \leq 3(m+1)^{-m}$ for all $m \geq 0$,

$$
\begin{aligned}
C_{is} &\leq \binom{i+s}{s}\binom{k-i-1}{k-i-s} \\
&\leq \left(\frac{e(i+s)}{s}\right)^{s}\left(\frac{k-i}{e}\right)^{k-i}\left(\frac{k-i-s}{e}\right)^{s-(k-i)}\left(\frac{e}{s-1}\right)^{(s-1)} \\
&\leq 3^{16}\left(\frac{ek}{s^2}\right)^{s}(k-i)^{k-i+1}(k-i-s+15)^{s-(k-i)}. \qquad (21)
\end{aligned}
$$

Similarly to case (b2), but without the constraint on the size of the longest disjoint interval, the extremal configuration for given s is now attained if we have $s-1$ disoint intervals of unit size and one interval containing all the other disjoint vertices. Thus

$$
\begin{aligned}
\frac{E_{is}}{R_k} &\leq \frac{\psi(0)!^{i-(s-1)}}{(n-2t)!}\,\psi(k-i-s+1)!\,\psi(1)!^{s-1} \\
&\leq \frac{3\psi(0)!^{i}}{(n-2t)!}\left(\frac{(k-i-s+15)r}{e}\right)^{\psi(k-i-s+1)+1/2}(4r)^{(s-1)r}.
\end{aligned}
$$

The second line follows by Stirling's formula (8). By the definition of h, the fact that $(k-i-s+15) \leq 3^3(k-i)$, and inequality (18), we now have

$$
\begin{aligned}
\frac{E_{is}}{R_k\,p^i} &\leq \frac{3^4}{d\sqrt{k}}(k-i)^{3/2}(k+1)^r \times \\
&\quad\times\left(\frac{e}{(k+1)r}\right)^{\psi(k-i)}\left(\frac{(k-i-s+15)r}{e}\right)^{\psi(k-i)-(s-1)r}(5r)^{(s-1)r} \\
&\leq \frac{3^7}{d\sqrt{k}}(k-i)^{3/2}\left(\frac{k-i-s+15}{(k+1)^{1-1/(k-i+1)}}\right)^{\psi(k-i)}\left(\frac{5e}{k-i-s+15}\right)^{(s-1)r} \\
&\leq \frac{3^{10}}{d\sqrt{k}}(k-i)^{5/2}\left(\frac{k-i-s+15}{k}\right)^{k-i}\left(\frac{5e}{k-i-s+15}\right)^{s-1}. \qquad (22)
\end{aligned}
$$

The last line follows since one may check that both of the bracketed terms in the second line are smaller than one. Combining inequalities (21) and (22), we have

$$
\delta_{is} \leq \frac{3^{30}}{d\sqrt{k}}(k-i)^4\left(\frac{3e^2k}{s^2}\right)^{s}\left(\frac{k-i}{k}\right)^{k-i}.
$$

It is easily verified that for fixed k and i, the expression $\phi(s) = \left(3e^2k/s^2\right)^s$ is maximized at $s = \sqrt{3k}$ if $i \leq k - \sqrt{3k}$ and is maximized at $s = k-i$

if $i \geq k - \sqrt{3k}$. Substituting the ensuing bounds on $\phi(s)$ and noting that $s \leq k - i \leq k$, a straightforward calculation shows that

$$\sum_{(1-\rho)k \leq i < k} \delta_i \leq \frac{3^{30}}{d\sqrt{k}} \left(k^6 (\rho e^2)^{\sqrt{3k}} + K \right) \leq \frac{1}{dk^{1/4}},$$

where K is some fixed constant.

Summing up the bounds in (a)–(c) now gives the desired result:

$$\frac{\Delta}{\mathbb{E}(X_{\mathcal{S}})^2} \leq \frac{8}{d} + k^2 \frac{1}{dk^3} + \frac{1}{dk^{1/4}} \leq \frac{9}{d}.$$

□

We remark that the above result is best possible up to a constant factor. This may be seen by noting that the contribution to $\Delta/\mathbb{E}(X_{\mathcal{S}})^2$ of those pairs of chains which have only the lowest element in common is asymptotically equal to $1/ed$.

In the previous section, we mentioned that the threshold in Theorem 1 could be tightened by a factor of $(n/k)^2$: this can be done by considering not just (x, y, r)-chains (where the level of the ith element x_i is fixed at $|x_i| = |x| + \lceil ir \rceil$), but allowing the level of x_i to deviate from this by a factor of $\sqrt{r}$. The calculations are then more cumbersome, but the techniques of this section still apply.

Finally, we state a slight variant of the previous lemma, which is needed in [11], the difference being that the regular chains under consideration are now "squeezed": let k be given as before and let $m = m(n)$ and $m' = m'(n)$ be given integer functions which are $o(n)$. In the definitions of r, ψ, and p in (5), (6), and (7) respectively, replace the term $2t$ by $m + m'$. Then we say that $(x_1 \subset \cdots \subset x_k)$ is an $[m, m', r]$-chain if $|x_i| = m + \lceil ir \rceil$ for all i.

Lemma 8 *Let $\mathcal{S}$ be the set of all $[m, m', r]$-chains in $\mathcal{P}(n)$. Then for large enough n, and p as above, we have*

$$\Delta(\mathcal{S}) \leq \frac{9}{d}\, \mathbb{E}(X_{\mathcal{S}})^2.$$

The proof of Lemma 7 goes through as before: we may first of all assume without loss of generality that $m \geq m'$; it turns out that the only difference arises when bounding the number of extensions: in the extremal configuration the largest disjoint interval is now situated at the bottom, which contributes an extra factor of at most $(\psi(l) + m)_m$ to the numerator of the E_{ils}/R_k. Similarly, the second largest interval contributes an extra factor of at most $(\psi(l) + m')_{m'}$. This is compensated by the additional $(n)_{m+m'}$ in the denominator. In fact the only reason for stating Lemma 8 separately was to avoid additional notation in the previous lemma and its proof.

We further remark that the above bound on $\Delta(\mathcal{S})$ is only required in [11]. For the purposes of this paper, it would have been slightly more convenient to bound $\mathbb{E}(X_{\mathcal{S}}^2)$ directly instead of $\Delta(\mathcal{S})$.

4 Proof of Theorem 1

The technically difficult part in the proof of Theorem 1 has been dealt with in Lemma 7, and hence this theorem may now be proved easily.

Proof of Theorem 1. For the first part assume that $p\,(k+1)^{n/k} \to 0$. Let Y_k be the number of k-element chains in $\mathcal{P}(n,p)$. The number of k-element chains in $\mathcal{P}(n)$ is at most $(k+1)^n$. To see this, note that we may associate with each k-element chain α a $(k+1)$-colouring of $\{1,\dots,n\}$ as follows. Suppose $\alpha = (x_1 \subset x_2 \cdots \subset x_k)$. Assign colour 1 to all elements in x_1, colour j to all elements in $x_j \setminus x_{j-1}$ for $1 < j \leq k$ and colour $k+1$ to $[n] \setminus x_k$. Hence

$$\mathbb{P}(Y_k > 0) \leq \mathbb{E}(Y_k) \leq (k+1)^n\, p^k = o(1),$$

which implies the first statement.

To prove the second statement, suppose $p\,(k+1)^{n/k}(k/n)^2 \to \infty$. Let r be given by (5). It is easily checked that this implies that p satisfies (7) with $d = d(n) \to \infty$. Now let $\mathcal{S}$ be the set of $(\emptyset,[n],r)$-chains, so that $X_{\mathcal{S}}$ is the number of $(\emptyset,[n],r)$-chains in $\mathcal{P}(n,p)$. Since $\mathrm{Var}(X_{\mathcal{S}}) \leq \Delta(\mathcal{S}) + \mathbb{E}(X_{\mathcal{S}})$ (see e.g. [1]), the desired result follows immediately from Chebyshev's inequality after applying Proposition 6 and Lemma 7 with $t = 0$. $\square$

5 Proof of Theorem 4

We shall use some results on branching processes in the proof of Theorem 4 to "amplify" the results of the previous section. This is a natural approach also exploited in [7]. We first present a few basic concepts and some notation.

Let X be a nonnegative integer valued random variable. A *Galton-Watson branching process* with family size X is a discrete Markov chain $(X_t)_{t=0}^{\infty}$ with the following properties: at time $t = 0$ we have one "ancestor", so we set $X_0 = 1$. For $t \geq 1$ define X_t to be the sum of X_{t-1} independent random variables, each having the same distribution as X (see e.g. [8] for details). As we shall only be concerned with the case when X has Poisson distribution with parameter c, in what follows $(X_t)_{t=0}^{\infty}$ will always refer to a branching process with this property. Let η be the *extinction probability* of this process:

$$\eta = \eta(c) = \mathbb{P}(X_t = 0 \text{ for some } t). \tag{23}$$

A simple result in branching process theory says that if $c > 1$ then η is the (unique) root of $s = \mathbb{E}(s^X) = e^{c(s-1)}$ in $0 < s < 1$. Another basic result (see [8] again) is that for any constant $k > 0$, we have

$$\lim_{t\to\infty} \mathbb{P}(X_t = k) = 0. \tag{24}$$

We now define branching processes in $\mathcal{P}(n)$. Consider any $x \in \mathcal{P}(n)$ at level m, say. An *upward* x-process $(Z_t^n)_{t=0}^{n-m} = (Z_t^n(x))_{t=0}^{n-m}$ is defined by letting Z_t^n denote the number of $z \in P_{m+t}(n,p)$ that are connected to x via an $(x,z,1)$-chain in $\mathcal{P}(n,p)$. A downward y-process is defined similarly. For simplicity, we will henceforth call an $(x,y,1)$-chain an xy-chain.

The following proposition will help us determine the number of descendants we will need at each end to link the "top" and "bottom" of $\mathcal{P}(n,p)$:

Proposition 9 *For any given $c \geq e$ define $\eta = \eta(c)$ as in (23). Given $0 < \varepsilon < 1-\eta$, there exists $a = a(c,\varepsilon) \in \mathbb{N}$ with $(43/44)^a \leq \varepsilon$ satisfying the following: for a random variable $Z \sim \mathrm{Bin}(\lceil 2a/(1-\eta-\varepsilon)\rceil, 1-\eta-\varepsilon)$, we have $\mathbb{P}(Z < a) < \varepsilon$.*

Proof. Certainly $\mathbb{P}(Z > a) \leq \mathbb{P}(|Z - \mathbb{E}(Z)| \geq a)$. Thus the result follows from Chernoff's bound. □

For positive integers m and n with $n > 2m$, denote by $\mathcal{P}_{m\ldots n-m}(n,p)$ the family of all sets from $\mathcal{P}(n,p)$ of cardinality at least m and at most $n-m$. Let $X \subseteq \mathcal{P}_m(n)$ and $Y \subseteq \mathcal{P}_{n-m}(n)$ be given. We say that X and Y *can be linked* if there are $x \in X$ and $y \in Y$ such that there is an xy-chain contained in $\mathcal{P}_{m+1\ldots n-m-1}(n,p)$.

Lemma 10 *Let $\varepsilon > 0$, $c \geq e$ and $m \in \mathbb{N}$ be given. Let $p = c/n$ and define $\eta = \eta(c)$ as in (23). Let $a = a(c,\varepsilon)$ be as in Proposition 9. Let*

$$b = \lceil 2a/(1-\eta-\varepsilon)\rceil .$$

Furthermore let $X \subseteq \mathcal{P}_m(n)$ with $|X| = b$ and $Y \subseteq \mathcal{P}_{n-m}(n)$ with $|Y| = b$ and $x \subset y$ for any $x \in X$ and $y \in Y$ be given. Then for large enough n

$$\mathbb{P}(X \text{ and } Y \text{ can be linked}) \geq 1 - \varepsilon.$$

Lemma 7 implies that a single pair $x \subset y$ is linked with probability bounded away from zero. Hence, to prove Lemma 10, we would like to consider several pairs at the same time to ensure that with arbitrarily high probability we can find at least one that is linked by a chain. The following lemma tells us that this is indeed possible, provided the elements at each level are far enough apart.

Given x, $y \in \mathcal{P}(n)$ with $x \subset y$, denote by $\mathcal{P}^{xy}$ the subset of $\mathcal{P}(n)$ consisting of all elements (including x and y) which are both subset of y and superset of x.

Lemma 11 *Let t, $h \in \mathbb{N}$ be given. Suppose that x_1, $x_2 \in \mathcal{P}_t(n)$ and y_1, $y_2 \in \mathcal{P}_{n-t}(n)$ are such that $x_1 \subset y_1$, $x_2 \subset y_2$, $d_H(x_1,x_2) \geq h$ and $d_H(y_1,y_2) \geq h$. Then the proportion of all x_1y_1-chains that meet $\mathcal{P}^{x_2y_2}$ is at most $2e^{-h/2}$, provided n is large enough.*

Proof. Let $\bar{x} = x_1 \cup x_2$ and $\bar{y} = y_1 \cap y_2$. Then one may easily verify that $\mathcal{P}^{x_1y_1} \cap \mathcal{P}^{x_2y_2} = \mathcal{P}^{\bar{x}\bar{y}}$. Furthermore $\bar{x}$ has cardinality at least $t+h$ and $\bar{y}$ has cardinality at most $n-(t+h)$. The crucial observation now is that any x_1y_1-chain which meets $\mathcal{P}^{x_2y_2}$ must be *contained* in $\mathcal{P}^{x_1\bar{y}} \cup \mathcal{P}^{\bar{x}y_1}$. Indeed, suppose an x_1y_1-chain contains a $z \in \mathcal{P}^{x_2y_2}$. Then all elements in that chain below z are contained in $\mathcal{P}^{x_1\bar{y}}$ and all elements above z are contained in $\mathcal{P}^{\bar{x}y_1}$. It follows that any x_1y_1-chain which meets $\mathcal{P}^{x_2y_2}$ intersects $(\mathcal{P}^{x_1\bar{y}} \cup \mathcal{P}^{\bar{x}y_1}) \cap \mathcal{P}_{\lfloor n/2\rfloor}(n)$. However, the fraction of such chains among all x_1y_1-chains is at most

$$\frac{|(\mathcal{P}^{x_1\bar{y}} \cup \mathcal{P}^{\bar{x}y_1}) \cap \mathcal{P}_{\lfloor n/2\rfloor}(n)|}{|\mathcal{P}^{x_1y_1} \cap \mathcal{P}_{n/2}(n)|} \leq 2\binom{n-2t-h}{\lfloor (n-2t)/2\rfloor} \bigg/ \binom{n-2t}{\lfloor (n-2t)/2\rfloor} \leq 2e^{-h/2}$$

and the lemma follows. $\square$

We are now ready to prove the Lemma 10. In what follows, we will make use of the following form of the second moment inequality:

$$\mathbb{P}(X_{\mathcal{S}} = 0) \leq 1 - \frac{\mathbb{E}(X_{\mathcal{S}})^2}{\mathbb{E}(X_{\mathcal{S}}^2)}. \tag{25}$$

A proof of this elementary inequality may be found in [2].

Proof of Lemma 10. Let a be as in Proposition 9 and set $h = \log 4a$. For each $x \in X$ consider the upward x-branching process. For each such process the probability of survival during the next h steps is at least $1 - \eta - \varepsilon$. We may assume that n is large enough to ensure that the probability that these processes intersect is at most ε. Thus, by the definition of b in Lemma 10 and by Proposition 9, with probability at least $1-2\varepsilon$ there are a processes which survive until level $m+h$. Pick one descendant $x_i \in \mathcal{P}_{m+h}(n)$ of each, where $1 \leq i \leq a$. Let $X' = \{x_1, \dots, x_a\}$. Note that we may find $\bar{X} \subset X$, $\bar{X} = \{\bar{x}_1, \dots, \bar{x}_a\}$, so that there is an $\bar{x}_i x_i$-chain in $\mathcal{P}(n,p)$ for all i. Analogously, with probability $1-2\varepsilon$, we may find a set $Y' = \{y_1, \dots, y_a\}$ with $Y' \subset \mathcal{P}_{n-m-h}(n)$ and which is linked to Y in a similar way. Now assume that n is large enough to ensure that with probability at least $1-\varepsilon$ these branching processes have the following property: $(x_i \backslash \bar{x}_i) \cap (x_j \backslash \bar{x}_j) = \emptyset$ for all pairs $i \neq j$, and similarly for the y_i. The latter implies that $d_H(x_i, x_j) \geq 2h$ and $d_H(y_i, y_j) \geq 2h$ for $i \neq j$.

Now let $\mathcal{S}_i$ be the set of $x_i y_i$-chains and let $\mathcal{T}_i \subseteq \mathcal{S}_i$ be the set of $x_i y_i$-chains not meeting $\bigcup_{j \neq i} \mathcal{P}^{x_j y_j}$. We let $Y_i = X_{\mathcal{S}_i}$ and $Y_i' = X_{\mathcal{T}_i}$, where the notation is as in (9). We also let $\Delta_i = \Delta(\mathcal{S}_i)$ and $\Delta_i' = \Delta(\mathcal{T}_i)$, using the notation of (10). Note that Proposition 6 certainly implies that $\mathbb{E}(Y_i) \leq \mathbb{E}(Y_i)^2$. Thus by the definition of $\Delta_{\mathcal{S}_i}$ and Lemma 7 (with $t = m+h$ and $d = 1$) we have

$$\mathbb{E}(Y_i^2) \leq \Delta_i + \mathbb{E}(Y_i) + \mathbb{E}(Y_i)^2 \leq 11\mathbb{E}(Y_i)^2.$$

Now $Y_i' \leq Y_i$ implies that $\mathbb{E}(Y_i'^2) \leq \mathbb{E}(Y_i^2)$. But since $2a\exp(-h) \leq 1/2$, Lemma 11 shows that $\mathbb{E}(Y_i') \geq \mathbb{E}(Y_i)/2$. Altogether this implies that

$$\frac{\mathbb{E}(Y_i')^2}{\mathbb{E}(Y_i'^2)} \geq \frac{\mathbb{E}(Y_i^2)}{4\mathbb{E}(Y_i^2)} \geq \frac{1}{44}.$$

Thus applying in turn the independence of the Y_i', the second moment inequality (25), and the definition of a in Proposition 9, we have (for n sufficiently large)

$$\mathbb{P}\left(\bigcap_{i=1}^{a} Y_i' = 0\right) = \prod_{i=1}^{a} \mathbb{P}\left(Y_i' = 0\right) \leq \left(\frac{43}{44}\right)^a \leq \varepsilon$$

and the lemma follows. $\square$

In the remainder of this section we shall define $(X_t^n)_{t=0}^n$ as the Markov chain of an upward $\emptyset$-process and $(Y_t^n)_{t=0}^n$ as that of a downward $[n]$-process. We need a final preparation for the proof of Theorem 4:

Proposition 12 *For a given $c \geq e$ define $\eta = \eta(c)$ as in (23) and let $p = c/n$. Let $0 < \varepsilon < 1-\eta$ and $b \in \mathbb{N}$ be given. Then there exist m_0, $n_0 \in \mathbb{N}$ so that for $m \in \{m_0, m_0+1\}$ and all $n \geq n_0$, we have*

$$\begin{aligned} \mathbb{P}(X_m^n = 0) &\geq \eta - \varepsilon, & (26)\\ \mathbb{P}(X_m^n \geq b) &\geq 1 - \eta - \varepsilon. & (27) \end{aligned}$$

Proof. Let $(X_m)_{m=0}^{\infty}$ be a Poisson branching process with parameter c. Note that $\mathbb{P}(X_m = 0) \uparrow \eta$. Thus by (24) there is an m_0 such that for $m \in \{m_0, m_0+1\}$ we have

$$\begin{aligned} \mathbb{P}(X_m = 0) &\geq \eta - \varepsilon/2,\\ \mathbb{P}(X_m \leq b \mid X_m > 0) &\leq \varepsilon/2. \end{aligned}$$

It is straightforward to show that for fixed m, the random variables X_m^n tend to X_m in distribution as $n \to \infty$. This implies that (26) and (27) hold for sufficiently large n. □

Proof of Theorem 4. Let $\varepsilon > 0$ be given. We now fix an $a = a(c,\varepsilon)$ satisfying the conditions in Proposition 9 and let $b = \lceil 2a/(1-\eta-\varepsilon) \rceil$. Then we fix $m_0 = m_0(b,\varepsilon)$ as in Proposition 12 and let $m = m_0 + 1$. We set n large enough to satisfy all the requirements below. Consider the branching processes $(X_t^n)_{t=0}^n$ and $(Y_t^n)_{t=0}^n$ defined earlier. We will distinguish several cases according to their (non)-extinction during the first m generations. Note that almost surely $\mathcal{P}(n,p)$ contains neither the bottom element $\emptyset$ nor the top element $[n]$. For the remainder of the proof we will condition on this event. Thus a necessary condition for the length of $\mathcal{P}(n,p)$ to equal $n-2$ is that $X_m^n > 0$ and $Y_m^n > 0$. By Proposition 12 this happens with probability $(1-\eta)^2 + \mathcal{O}(\varepsilon)$. But the same proposition also tells us that in this case with probability $1-\varepsilon$ we have sets $X \subseteq \mathcal{P}_m(n,p)$ and $Y \subseteq \mathcal{P}_{n-m}(n,p)$ with $|X| = |Y| = b$ whose elements are the descendants of the respective branching processes. As m and b are fixed, with probability $1-\varepsilon$ each element of X will be contained in every element of Y. Since the event that X and Y can be linked depends only on $\mathcal{P}_{m+1\ldots n-m-1}(n,p)$, Lemma 10 then implies that

$$\mathbb{P}(\text{length}\,\mathcal{P}(n,p) = n-2) = (1-\eta)^2 + \mathcal{O}(\varepsilon).$$

Now consider the case when $X_m^n = 0$ and $Y_m^n > 0$, which by Proposition 12 occurs with probability $\eta(1-\eta) + \mathcal{O}(\varepsilon)$. Since we are assuming that $\emptyset$, $[n] \notin \mathcal{P}(n,p)$, in this case the length of $\mathcal{P}(n,p)$ is at most $n-3$. Let k satisfy $(\eta+\varepsilon)^k \leq \varepsilon$ and let $x_i = \{i\}$ for $i = 1,\ldots,k$. For each x_i, consider the upward x_i-branching process. Note that with probability $1-\varepsilon$ these branching processes are disjoint up to level m, i.e., $x_i' \neq x_j'$ whenever $i \neq j$, $x_i \subset x_i'$, $x_j \subset x_j'$, and x_i', $x_j' \in \mathcal{P}_{1\ldots m}(n,p)$. Another application of Proposition 12 tells us that with probability $1-\varepsilon$ at least one of the branching processes has (at least) b descendants in $\mathcal{P}_m(n,p)$. This time we pick any $X' \subseteq \mathcal{P}_m(n,p)$ with $|X'| = b$ from the set of descendants of this branching process and we pick any $Y \subseteq \mathcal{P}_{n-m}(n,p)$ as above. Again it is easy to see that with probability close to one, X' and Y satisfy the requirements of Lemma 10 and thus in this case $\mathcal{P}(n,p)$ has length exactly $n-3$. By symmetry the same applies if $X^n_m > 0$ and $Y^n_m = 0$ and thus, putting all of the above together, we have

$$\mathbb{P}(\text{length}\,\mathcal{P}(n,p) = n-3) = 2\eta(1-\eta) + \mathcal{O}(\varepsilon).$$

Finally suppose that $X^n_m = 0$ and $Y^n_m = 0$, which occurs with probability $\eta^2 + \mathcal{O}(\varepsilon)$. As above, with sufficiently high probability one can find suitable $X' \subseteq \mathcal{P}_m(n,p)$ and $Y' \subseteq \mathcal{P}_{n-m}(n,p)$ by considering k branching processes starting in $\mathcal{P}_1(n,p)$ and $\mathcal{P}_{n-1}(n,p)$, respectively, so that in this case Lemma 10 implies that $\mathcal{P}(n,p)$ has length $n-4$. $\square$

We remark that it is easy to strengthen Corollary 2 by replacing the condition $pn^r \to 0$ by $pn^r \leq c$ for some c depending only on the constant r. The methods of this section would also enable us to replace the condition $pn^r \to \infty$ by $pn^r \geq C$ for some C depending only on r. However, the analysis is rather cumbersome if r is not an integer, so we have chosen to omit it.

References

[1] N. Alon and J.H. Spencer, *The probabilistic method*, John Wiley & Sons, New York, 1992.

[2] B. Bollobás, *Random graphs*, Academic Press, London, 1985.

[3] B. Bollobás and G. Brightwell, Random high-dimensional orders, *Adv. Appl. Prob.*, **27**, 161–184 (1995).

[4] B. Bollobás, Y. Kohayakawa, and T. Łuczak, On the diameter and radius of random subgraphs of the cube, *Random Struct. Alg.*, **5**, 627–648 (1994).

[5] B. Bollobás, Y. Kohayakawa, and T. Łuczak, Connectivity properties of random subgraphs of the cube, *Random Struct. Alg.*, **6**, 221–230 (1995).

[6] G. Brightwell, Models of random partial orders, *Surveys in Combinatorics 1993* (K. Walker, Ed.), pp. 53–83, Cambridge Univ. Press, Cambridge, UK.

[7] J.A. Fill and R. Pemantle, Percolation, first-passage percolation and covering times for Richardson's model of the n-cube, *Annals Appl. Probab.*, **3** (2), 593–629 (1993).

[8] T.E. Harris, *The theory of branching processes*, Springer Verlag, Berlin, 1963.

[9] Y. Kohayakawa and B. Kreuter, The width of random subsets of Boolean lattices, submitted (1998).

[10] B. Kreuter, Small sublattices in random subsets of Boolean lattices, *Random Struct. Alg.*, **13**, 383–407 (1998).

[11] D. Osthus, Maximum antichains in random subsets of a finite set, *Journal of Comb. Theory A*, to appear.

[12] A. Rényi, On random subsets of a finite set, *Mathematica (Cluj)*, **3** (26), 355–362 (1961).