Maximum Antichains in Random Subsets of a Finite Set

to appear in *Journal of Combinatorial theory, Series A*

Deryk Osthus
Institut für Informatik, Humboldt-Universität zu Berlin
Unter den Linden 6, 10099 Berlin, Germany
E-mail address: osthus@informatik.hu-berlin.de

Abstract

We consider the random poset $\mathcal{P}(n,p)$ which is generated by first selecting each subset of $[n]=\{1,\dots,n\}$ with probability p and then ordering the selected subsets by inclusion. We give asymptotic estimates of the size of the maximum antichain for arbitrary $p=p(n)$. In particular, we prove that if $pn/\log n \to \infty$, an analogue of Sperner's theorem holds: almost surely the maximum antichain is (to first order) no larger than the antichain which is obtained by selecting all elements of $\mathcal{P}(n,p)$ with cardinality $\lfloor n/2 \rfloor$. This is almost surely not the case if $pn \not\to \infty$.

1 Introduction and results

Let $\mathcal{P}(n)$ be the power set of $[n]=\{1,\dots,n\}$. Call a subset $A \subseteq \mathcal{P}(n)$ an *antichain* if A contains no pair x, y so that $x \subset y$ and $x \neq y$. A classical result of Sperner says that a maximum antichain may be obtained by picking all subsets of $\mathcal{P}(n)$ which have cardinality $\lfloor n/2 \rfloor$. The main result of this paper is that a similar phenomenon occurs when one considers random subsets of $\mathcal{P}(n)$.

We obtain such a random subset $\mathcal{P}(n,p)$ by selecting each element of $\mathcal{P}(n)$ with probability p independently of all other subsets. If we order the elements of $\mathcal{P}(n)$ by inclusion, we may regard $\mathcal{P}(n,p)$ as a random partially ordered set. This model was first considered by Rényi [7], who obtained the threshold for the property that $\mathcal{P}(n,p)$ is not an antichain itself.

A closely related problem is that of determining the *width* of $\mathcal{P}(n,p)$ for given $p=p(n)$, where the width of $\mathcal{P}(n,p)$ is defined to be the cardinality of the largest antichain contained in $\mathcal{P}(n,p)$. We write $\mathcal{P}_i(n,p)$ for the set of all those elements of $\mathcal{P}(n,p)$ which have cardinality i and call it the *ith level* of $\mathcal{P}(n,p)$. Note that the expected size of the ith level is just $p\binom{n}{i}$. We write *a.s.* to mean "with probability tending to one as n tends to infinity".

Theorem 1 *If $pn/\log n \to \infty$, then* a.s. *we have*

$$\operatorname{width} \mathcal{P}(n,p) = (1+o(1))\, p \binom{n}{\lfloor n/2 \rfloor}.$$

An application of Chernoff's bound immediately yields that the size of the middle level of $\mathcal{P}(n,p)$ is *a.s.* sharply concentrated about its expected value $p\binom{n}{\lfloor n/2 \rfloor}$. Thus Theorem 1 is indeed a probabilistic analogue of Sperner's theorem: for large enough p, one may obtain an antichain whose size is *a.s.* within a factor $1+o(1)$ of the width simply by picking all elements of $\mathcal{P}(n,p)$ which have cardinality $\lfloor n/2 \rfloor$.

Theorem 1 is a special case (corresponding to $r = s = 1$) of the following general theorem, which gives upper and lower bounds on the width of $\mathcal{P}(n,p)$ for any given p.

Theorem 2 *Given $p = p(n)$, let $r = r(n)$ be an integer function which is large enough to satisfy*

$$p \left(\frac{n}{r}\right)^r / \log n \to \infty \tag{1}$$

Then a.s. *we have*

$$\operatorname{width} \mathcal{P}(n,p) \leq (1+o(1))\, p \sum_{-r/2 \leq j < r/2} \binom{n}{\lfloor n/2 \rfloor + j}.$$

If $p \to 0$, let $s = s(n) \geq 1$ be an integer function which is small enough to satisfy

$$p \left(\frac{n}{s}\right)^{s-1} \to 0. \tag{2}$$

If $p \not\to 0$, let $s = 1$. Then a.s. *we have*

$$\operatorname{width} \mathcal{P}(n,p) \geq (1+o(1))\, p \sum_{-s/2 \leq j < s/2} \binom{n}{\lfloor n/2 \rfloor + j}.$$

The upper bound in Theorem 2 is obtained by a double counting argument applied to chains in $\mathcal{P}(n,p)$, following the elegant proof of Sperner's theorem by Lubell [6, 1]. Here a *chain* is a subset of $\mathcal{P}(n)$ whose elements are pairwise comparable. To make this argument work, we make heavy use of a technical result from [4]: condition (1) essentially guarantees the existence of many chains in $\mathcal{P}(n,p)$ whose cardinalities are close to n/r.

The lower bound is obtained simply by showing that the expected number of relations between the elements contained in

$$A' = \bigcup_{-\overline{s}/2 \leq j < \overline{s}/2} \mathcal{P}_{\lfloor n/2 \rfloor + j}(n,p)$$

is small, where $\overline{s} = (1+o(1))s$. This implies that for given p we may always find an antichain A in $\mathcal{P}(n,p)$ which has at least $(1+o(1))|A'|$ elements, by discarding from A' any elements which are contained in another element of A'.

How good are the bounds in Theorem 2? Suppose r is a minimal integer function satisfying (1) in the sense that r satisfies (1) but $r-1$ does not. Similarly suppose s is a maximal integer function satisfying (2) in the sense that s satisfies (2) but $s+1$ does not. Then it is easily checked that $0 \leq r - s \leq 1$, and thus the width of $\mathcal{P}(n,p)$ is determined up to a factor of $1 + 1/r$. If $r \to \infty$, or equivalently, if p decays faster than any polynomial in n, this gives matching (up to first order) upper and lower bounds on the width of $\mathcal{P}(n,p)$:

Corollary 3 *Given $p = p(n)$, let $r = r(n)$ be defined by $p = (r/n)^r$. Then* a.s. *we have*

$$\operatorname{width} \mathcal{P}(n,p) = (1 + \mathcal{O}(1/r) + o(1)) \; p \sum_{-r/2 \le j < r/2} \binom{n}{\lfloor n/2 \rfloor + j}.$$

Corollary 3 was also obtained by Kohayakawa and Kreuter [3] if r is of the form $r = c\sqrt{n}$, for any constant $c > 0$. Observe that this is an especially interesting period in the evolution of $\mathcal{P}(n,p)$: Corollary 3 implies that

$$\frac{\operatorname{width} \mathcal{P}(n,p)}{|\mathcal{P}(n,p)|} \to \quad \begin{cases} 0 & \text{if } r/\sqrt{n} \to 0 \\ 1 & \text{if } r/\sqrt{n} \to \infty. \end{cases} \tag{3}$$

If $r = o(\sqrt{n})$ however, the matching argument used in [3] provides no further information beyond (3) about the width of $\mathcal{P}(n,p)$.

Finally, we point out that it would be very interesting to know whether the condition $pn/\log n \to \infty$ in Theorem 1 could be replaced by $pn \to \infty$. The following proposition shows that the latter condition is certainly necessary (similar remarks hold also for the factor $\log n$ in Theorem 2).

Proposition 4 *Suppose $p = c/n$ where $c > 0$ is a fixed constant. Then*

$$\operatorname{width} \mathcal{P}(n,p) \ge (1 + o(1))\,(1 + e^{-c/2})\, p \binom{n}{\lfloor n/2 \rfloor}.$$

Proof. Consider the bipartite graph whose vertex classes are $V_1 = \mathcal{P}_{\lfloor n/2 \rfloor}(n,p)$ and $V_2 = \mathcal{P}_{\lfloor n/2 \rfloor + 1}(n,p)$, with an edge between $v_1 \in V_1$ and $v_2 \in V_2$ if $v_1 \subset v_2$. The probability that a vertex $v_2 \in V_2$ is isolated is equal to $(1-p)^{\lfloor n/2 \rfloor + 1}$. Let Z be the number of isolated vertices in V_2. Then

$$\mathbb{E}[Z] = (1 + o(1)) e^{-c/2} \mathbb{E}[|V_2|] = (1 + o(1)) e^{-c/2} p \binom{n}{\lfloor n/2 \rfloor}.$$

Using Chebyshev's inequality, it is easily verified that *a.s.* $Z = (1 + o(1))\mathbb{E}[Z]$. The union of V_1 and the isolated vertices in V_2 is thus *a.s.* an antichain whose size satisfies the requirements. □

The remainder of this note is organized as follows: in Section 2 we derive properties of "regular" chains in $\mathcal{P}(n,p)$, which will enable us to prove Theorem 2 in Section 3. We refer the reader to Kreuter [5] for further results on the evolution of $\mathcal{P}(n,p)$ and to Brightwell [2] for a survey on other models of random posets.

2 The chain cover

To prove the upper bound in Theorem 2 we need a covering of $\mathcal{P}(n,p)$ with chains which is well behaved in the sense that for most elements $a \in \mathcal{P}(n,p)$, the number of chains meeting a is large and very close to its expected value. It turns out that a cover whose chain elements are equally spaced and sufficiently far apart will do: we will quote a technical lemma from [4] which gives sufficient conditions on p and the average distance r between chain elements for the desired concentration result to hold. We shall then deduce two corollaries which will be of crucial importance in the proof of the main theorem.

Before we can state this technical lemma, we need some notation. Let $\mathcal{S}$ be a family of subsets of $\mathcal{P}(n)$. For $\alpha \in \mathcal{S}$, let $I_\alpha = I_\alpha(\mathcal{S}) = 1$ if $\alpha \subseteq \mathcal{P}(n,p)$ and $I_\alpha = I_\alpha(\mathcal{S}) = 0$ otherwise. Now let

$$X_{\mathcal{S}} = \sum_{\alpha \in \mathcal{S}} I_\alpha. \tag{4}$$

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

$$\Delta(\mathcal{S}) = \sum_{\alpha \sim \beta} \mathbb{E}\left[I_\alpha I_\beta\right], \tag{5}$$

where the sum is over all ordered pairs $(\alpha, \beta) \in \mathcal{S}$ with $\alpha \sim \beta$. Now let $m = o(n)$, $\overline{m} = o(n)$ and $k < n-1$ be given integer functions of n. Let

$$r = \frac{n - m - \overline{m} - 1}{k+1}. \tag{6}$$

Then we say that $(x_1 \subset \cdots \subset x_k)$ is an $[m, \overline{m}, r]$-chain if $|x_i| = m + \lceil ir \rceil$ for all i.

Lemma 5 [4] *Let r be as in* (6) *and suppose that p satisfies*

$$p(n) \geq d\,(r+1)! \left(\frac{e}{(n - m - \overline{m})(k+1)^{1/k}} \right)^r \tag{7}$$

with $d = d(n) \geq 1$. Let $\mathcal{S}$ be the set of all $[m, \overline{m}, r]$-chains in $\mathcal{P}(n)$. Then, for large enough n, we have

$$\Delta(\mathcal{S}) \leq \frac{9}{d} \mathbb{E}\left[X_{\mathcal{S}}\right]^2. \tag{8}$$

Now we are ready to define a chain cover of $\mathcal{P}(n)$ and use Lemma 5 to show it has the desired properties. Given integer functions $r = r(n)$ and $j = j(n)$ so that r is $o(n)$ and $j < r$, let $k = \lfloor n/r \rfloor - 2$. We remark that it is straightforward to check that this is consistent with (6), setting $m = j$ and $\overline{m} = n - 1 - (\lfloor n/r \rfloor - 1)r - j$. We say that $(x_1 \subset \cdots \subset x_k)$ is a $[j, r]$-chain if $|x_i| = j + ir$ for all i. Lemma 5 then immediately implies the following:

Corollary 6 *Let $\varepsilon > 0$ be given. Suppose that*

$$p(n) \left(\frac{n}{6r} \right)^r / \log n \to \infty. \tag{9}$$

Let $\mathcal{S}$ be the set of $[j, r]$-chains in $\mathcal{P}(n)$. Then

$$\mathbb{P}\left[X_{\mathcal{S}} \geq (1+\varepsilon)\mathbb{E}[X_{\mathcal{S}}]\right] = o(1/r).$$

Proof. It is easily checked that if p satisfies (9), then it also satisfies (7) with $d = r\log n$, say. Another a straightforward application of Stirling's formula yields that, for large enough n,

$$\begin{aligned}\mathbb{E}[X_{\mathcal{S}}] &= \binom{n}{j+r, r, \dots, r, n-kr-j} p^k \\ &\geq \left(\frac{n}{e}\right)^{n-kr} \frac{d}{(j+r)_j(n-kr-j)!} \geq d. \end{aligned} \tag{10}$$

The result now follows immediately from Chebyshev's inequality using the fact that $\mathrm{Var}[X_{\mathcal{S}}] \leq \Delta(\mathcal{S}) + \mathbb{E}[X_{\mathcal{S}}]$. For a proof of the latter inequality see e.g. [1]. □

To obtain our second corollary of Lemma 5, we need the following version of Janson's inequality (for a proof see e.g. [1] again): given $\tau > 0$,

$$\mathbb{P}[X_{\mathcal{S}} \leq (1-\tau)\mathbb{E}[X_{\mathcal{S}}]] \leq \exp\left(-\frac{\tau^2\mathbb{E}[X_{\mathcal{S}}]^2}{2(\mathbb{E}[X_{\mathcal{S}}] + \Delta(\mathcal{S}))}\right). \tag{11}$$

Let X_a denote the number of $[j,r]$-chains (for any $j < r$) which meet a and which are contained in $\mathcal{P}(n,p) \cup \{a\}$. Suppressing the dependence on r, let $\mathcal{P}^j(n)$ be the set of elements in $\mathcal{P}(n)$ which intersect some $[j,r]$-chain. In other words, $\mathcal{P}^j(n)$ consists of every rth level of $\mathcal{P}(n)$, starting with the $(j+r)$th.

Corollary 7 *Let $\varepsilon > 0$ be given and fix $a \in \mathcal{P}(n)$ with $n/3 \leq |a| \leq 2n/3$. Then if p satisfies* (9) *and n is sufficiently large, we have*

$$\mathbb{P}[X_a \leq (1-\varepsilon)\mathbb{E}[X_a]] \leq 2n^{-2}.$$

Proof. This time we apply Lemma 5 not to the whole of $\mathcal{P}(n)$ but to $\mathcal{P}(|a|)$ and $\mathcal{P}(n-|a|)$ separately, with

$$d(n) = \frac{80}{\varepsilon^2}\log n. \tag{12}$$

Consider $\mathcal{P}(|a|)$ first: fix j so that $a \in \mathcal{P}^j(n)$. Then let $m' = 0$, $\overline{m}' = j-1$, and $k' = \frac{|a|-j}{r} - 1$. Let $\mathcal{S}_1$ be the set of all chains $(x_1 \supset x_2 \supset \cdots \supset x_{k'})$ in $\mathcal{P}(|a|)$ with $|x_i| = |a| - ir$ for all i. Using the same notation as in (4), let $X_{\mathcal{S}_1}$ be the number of such chains in $\mathcal{P}(n,p)$. Similarly to (10) in the proof of the previous corollary, for large enough n we have

$$\mathbb{E}[X_{\mathcal{S}_1}] = \binom{|a|}{r, \dots, r, r+j} p^{k'} \geq d(n).$$

It is also easily checked that if $p(n)$ satisfies (9), then $p(|a|)$ satisfies (7) with n replaced by $|a|$; k, m, and $\overline{m}$ replaced by k', m', and $\overline{m}'$ respectively; but keeping $r(n)$ as before and $d(n)$ as in (12). Thus the conclusion (8) of Lemma 5 holds for the set $\mathcal{S}_1$ and $d(n)$ as in (12). Now apply Janson's inequality (11) to obtain

$$\mathbb{P}[X_{\mathcal{S}_1} \leq (1-\tau)\mathbb{E}[X_{\mathcal{S}_1}]] \leq e^{-\tau^2 d/20} = n^{-4\tau^2/\varepsilon^2}.$$

For $\mathcal{P}(n-|a|)$ we define $\mathcal{S}_2$ similarly and apply Janson's inequality in the same way. Note that $X_a = X_{\mathcal{S}_1}X_{\mathcal{S}_2}$. By independence we then have

$$\begin{aligned}\mathbb{P}\left[X_a \leq (1-\varepsilon)\mathbb{E}[X_a]\right] &= \mathbb{P}\left[X_{\mathcal{S}_1}X_{\mathcal{S}_2} \leq (1-\varepsilon)\mathbb{E}[X_{\mathcal{S}_1}]\mathbb{E}[X_{\mathcal{S}_2}]\right] \\ &\leq \sum_{i=1,2} \mathbb{P}\left[X_{\mathcal{S}_i} \leq \sqrt{1-\varepsilon}\,\mathbb{E}[X_{\mathcal{S}_i}]\right]\end{aligned}$$

Since $\sqrt{1-\varepsilon} \leq 1-\varepsilon/2$, the proof now follows by setting $\tau = \varepsilon/2$. $\square$

Finally, let $\mathcal{P}^j_{max}(n)$ be a level of $\mathcal{P}^j(n)$ which is closest to the middle and thus contains the most elements.

Proposition 8 *Suppose that for some j, we have $a \in \mathcal{P}^j(n)$ and furthermore that $a' \in \mathcal{P}^j_{max}(n)$. Then $\mathbb{E}[X_a] \geq \mathbb{E}[X_{a'}]$.*

Proof.

$$\mathbb{E}[X_a] = \binom{|a|}{j+r, r, \ldots, r}\binom{n-|a|}{r, \ldots, r, n-kr-j} p^{k-1},$$

and the result follows by the log-convexity of factorials. $\square$

3 Proof of the main theorem

With the results of the previous section at hand, the proof of the upper bound of Theorem 2 proceeds as follows: first we show that we may assume that $\mathcal{P}(n,p)$ does not contain many more "regular" chains than we would expect. Now suppose we have a large antichain A in $\mathcal{P}(n,p)$. Since A intersects any chain at most once, this means that A (and thus $\mathcal{P}(n,p)$) contains many "bad" elements in the sense that these elements lie in fewer "regular" chains than one would expect. But again this is extremely unlikely.

Proof of Theorem 2. We first prove the lower bound. The case $s=1$ is trivial, so suppose that p and s satisfy (2), where $p \to 0$. Since the assertion implies that *a.s.* width $\mathcal{P}(n,p) \geq (1+o(1))|\mathcal{P}(n,p)|$ whenever $s/\sqrt{n} \to \infty$, we may also assume that $s = o(n)$. Thus one may define $\overline{s} \leq s$ so that $\overline{s} = (1+o(1))s$ and

$$p\,\overline{s}^2 \left(\frac{en}{\overline{s}-1}\right)^{\overline{s}-1} \to 0. \qquad (13)$$

Now consider the expected number of edges $\mathbb{E}[|E(G)|]$ in the $\overline{s}$-partite graph G whose vertex classes are $V_j = \mathcal{P}_{\lfloor n/2\rfloor + j}(n,p)$ for all j satisfying $\overline{s}/2 \leq j < \overline{s}/2$, with an edge between $v_i \in V_i$ and $v_j \in V_j$ for $i \neq j$ if $v_i \subset v_j$ or $v_j \subset v_i$. Then

Stirling's formula and our assumption (13) on $\overline{s}$ imply that

$$\begin{aligned}
\mathbb{E}[|E(G)|] &= \sum_{-\overline{s}/2\le \ell<\ell'<\overline{s}/2} \binom{n}{\lfloor n/2\rfloor + l}\binom{\lfloor n/2\rfloor-\ell}{\ell'-\ell}p^2\\
&\le \sum_{-\overline{s}/2\le \ell<\ell'<\overline{s}/2} \binom{n}{\lfloor n/2\rfloor}\left(\frac{en}{\ell'-\ell}\right)^{\ell'-\ell}p^2\\
&\le p\binom{n}{\lfloor n/2\rfloor}p\,\overline{s}^2\left(\frac{en}{\overline{s}-1}\right)^{\overline{s}-1}\\
&= o\left(p\binom{n}{\lfloor n/2\rfloor}\right).
\end{aligned}$$

This proves the lower bound since it implies that an antichain of the required size may be obtained by picking any maximal independent set in G.

We now prove the upper bound. Suppose first that p and r satisfy (9), so that we may apply Corollaries 6 and 7. Furthermore note that we may assume that $r = \mathcal{O}(\sqrt{n})$. Indeed, if this is not the case, then a simple application of Stirling's formula tells us that in that case Theorem 2 just states that the width of $\mathcal{P}(n,p)$ is *a.s.* at most $(1+o(1))|\mathcal{P}(n,p)|$, which is of course trivial. Thus, defining $\mathcal{P}^j_{max}(n,p) = \mathcal{P}^j_{max}(n)\cap\mathcal{P}(n,p)$, another application of Stirling's formula implies that we may assume that

$$\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right] = p\,|\mathcal{P}^j_{max}(n)| \ge p\,2^n/n, \tag{14}$$

which will be convenient later on. For given $\varepsilon > 0$ define the event

$$\mathcal{A}_\varepsilon = \left\{\exists \text{ an antichain } A\subset\mathcal{P}(n,p) \text{ with } |A|\ge(1+\varepsilon)p\sum_{-r/2\le j<r/2}\binom{n}{\lfloor n/2\rfloor+j}\right\}.$$

For all j with $0\le j<r$ define the event

$$\mathcal{A}^j_\varepsilon = \left\{\mathcal{P}(n,p) \text{ contains an antichain } A \text{ with } |A^j|\ge(1+\varepsilon)\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right]\right\},$$

where A^j is defined by $A^j = A\cap\mathcal{P}^j(n)$. Given ε_1, $\nu>0$, our goal is to show that $\mathbb{P}[\mathcal{A}_{\varepsilon_1}]\le 3\nu$ for large enough n. Clearly, if $\mathcal{A}_{\varepsilon_1}$ holds, then we must have $\mathcal{A}^j_{\varepsilon_1}$ for some j. Thus,

$$\mathbb{P}[\mathcal{A}_{\varepsilon_1}] \le \mathbb{P}\left[\bigcup_j \mathcal{A}^j_{\varepsilon_1}\right] \le \sum_j \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\right]. \tag{15}$$

Fix some j and let X be equal to the number of $[j,r]$-chains contained in $\mathcal{P}(n,p)$. Set $\varepsilon_2 = \varepsilon_1/5$ and define the event

$$\mathcal{B}_{\varepsilon_2} = \{X < (1+\varepsilon_2)\mathbb{E}[X]\}.$$

Then for large enough n, Corollary 6 implies that

$$\begin{aligned}\mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\right] &= \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1} \cap \mathcal{B}_{\varepsilon_2}\right] + \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1} \cap \overline{\mathcal{B}}_{\varepsilon_2}\right] \\ &\le \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1} \cap \mathcal{B}_{\varepsilon_2}\right] + \mathbb{P}\left[\overline{\mathcal{B}}_{\varepsilon_2}\right] \\ &\le \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1} \cap \mathcal{B}_{\varepsilon_2}\right] + \nu/r.\end{aligned}$$

Now define the event

$$\begin{aligned}\mathcal{C}_{\varepsilon_1\varepsilon_2} = \Big\{\mathcal{P}(n,p) \text{ contains an antichain } A \text{ with } |A^j| \ge (1+\varepsilon_1)\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right] \\ \text{and } \sum_{a\in A^j} X_a < (1+\varepsilon_2)\mathbb{E}[X]\Big\}.\end{aligned}$$

Note that $\sum_{a\in A^j} X_a \le X$ implies that

$$\mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1} \cap \mathcal{B}_{\varepsilon_2}\right] \le \mathbb{P}\left[\mathcal{C}_{\varepsilon_1\varepsilon_2}\right].$$

For a given $\varepsilon > 0$ we say that $a \in \mathcal{P}(n)$ is *ε-bad* if $X_a \le (1-\varepsilon)\mathbb{E}[X_a]$. Setting $\mathcal{P}^j(n,p) = \mathcal{P}^j(n) \cap \mathcal{P}(n,p)$, we then let Y_ε be the number of ε-bad elements in $\mathcal{P}^j(n,p)$. Define

$$\mathcal{D}_{\varepsilon_2} = \left\{Y_{\varepsilon_2} \ge \varepsilon_2\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right]\right\}.$$

It is in the proof of the following claim that we will apply the double counting argument that Lubell used in his proof of Sperner's theorem.

Claim. We have $\mathcal{C}_{\varepsilon_1\varepsilon_2} \subset \mathcal{D}_{\varepsilon_2}$, and hence $\mathbb{P}\left[\mathcal{C}_{\varepsilon_1\varepsilon_2}\right] \le \mathbb{P}\left[\mathcal{D}_{\varepsilon_2}\right]$.

Proof. Suppose to the contrary that $\mathcal{C}_{\varepsilon_1\varepsilon_2} \cap \overline{\mathcal{D}}_{\varepsilon_2}$ holds and consider an antichain A guaranteed by $\mathcal{C}_{\varepsilon_1\varepsilon_2}$. Since we assumed that $\mathcal{D}_{\varepsilon_2}$ is false, it follows that

$$\left|\{a \in A^j : a \text{ is not } \varepsilon_2\text{-bad}\}\right| \ge (1+\varepsilon_1-\varepsilon_2)\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right].$$

The crucial point now is that, by double counting chains in $\mathcal{P}^j(n,p)$,

$$X = \sum_{a\in\mathcal{P}^j_{max}(n,p)} X_a,$$

and furthermore X_a is independent of the event that $a \in \mathcal{P}(n,p)$. Thus, by linearity of expectation, we have (for any $a' \in \mathcal{P}^j_{max}(n)$)

$$\mathbb{E}[X] = \mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right]\ \mathbb{E}[X_{a'}].$$

Combining all this with Proposition 8, we have that (again for any $a' \in \mathcal{P}^j_{max}(n)$)

$$\begin{aligned}\sum_{a\in A^j} X_a &\ge (1+\varepsilon_1-\varepsilon_2)\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right]\ (1-\varepsilon_2)\mathbb{E}[X_{a'}] \\ &= (1+\varepsilon_1-\varepsilon_2)(1-\varepsilon_2)\mathbb{E}[X] \\ &\ge (1+2\varepsilon_2)\mathbb{E}[X],\end{aligned}$$

since $\varepsilon_2 = \varepsilon_1/5$. But this contradicts the definition of $\mathcal{C}_{\varepsilon_1\varepsilon_2}$. $\square$

Thus it remains to show that $\mathbb{P}[\mathcal{D}_{\varepsilon_2}]$ is small. But Corollary 7 tells us that if $n/3 \le |a| \le 2n/3$, then

$$\mathbb{P}[a \in \mathcal{P}(n) \text{ is } \varepsilon_2\text{-bad}] \le 2n^{-2} \le \frac{\nu\varepsilon_2}{r} \frac{\mathbb{E}[|\mathcal{P}^j_{max}(n,p)|]}{p\,2^n}. \tag{16}$$

The last inequality follows from our assumption that $r = \mathcal{O}(\sqrt{n})$, which gives the lower bound (14) on $\mathbb{E}[|\mathcal{P}^j_{max}(n,p)|]$. Inequality (16) immediately yields an upper bound on the expected number of bad elements in $\mathcal{P}^j(n,p)$. Indeed, writing $\sum'_a$ for the sum over $a \in \mathcal{P}(n)$ with $n/3 \le |a| \le 2n/3$ and writing $\sum''_a$ for the sum over $a \in \mathcal{P}(n)$ with $|a| < n/3$ or $|a| > 2n/3$, we have

$$\begin{aligned}\mathbb{E}[Y_{\varepsilon_2}] &= \sum_{a\in\mathcal{P}^j(n)} p\,\mathbb{P}[a \text{ is } \varepsilon_2\text{-bad}] \le \sum_{a\in\mathcal{P}(n)} p\,\mathbb{P}[a \text{ is } \varepsilon_2\text{-bad}] \\ &\le {\sum}'_a p\,\mathbb{P}[a \text{ is } \varepsilon_2\text{-bad}] + {\sum}''_a p \\ &\le \frac{2\nu\varepsilon_2}{r}\,\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right].\end{aligned}$$

In the last line we made use of the fact that the bottom and top third of $\mathcal{P}(n)$ contain (very crudely) at most $2^n/n^4$ elements. Markov's inequality now implies that $\mathbb{P}[\mathcal{D}_{\varepsilon_2}] \le \mathbb{P}[Y_{\varepsilon_2} \ge \frac{r}{2\nu}\mathbb{E}[Y_{\varepsilon_2}]] \le 2\nu/r$, and thus

$$\mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\right] \le \mathbb{P}[\mathcal{C}_{\varepsilon_1\varepsilon_2}] + \nu/r \le \mathbb{P}[\mathcal{D}_{\varepsilon_2}] + \nu/r \le 3\nu/r,$$

as required. By (15), this completes the proof of Theorem 2 except for the fact that for the proof we assumed that p and r satisfy (9) instead of the original condition (1) in the statement of Theorem 2. This can be remedied immediately however, by noting that if p and r satisfy (1), then one may find an integer function $r' \ge r$ with $r' = (1+o(1))r$ so that p and r' satisfy (9). □

Note that the methods of Section 2 show that if $pn \to \infty$, then *a.s.* the fraction of elements in $\mathcal{P}(n,p)$ that are ε-bad tends to zero for any $\varepsilon > 0$. However, for our argument to work, we need the number of ε-bad elements in $\mathcal{P}(n,p)$ to be much less than the size of the middle level of $\mathcal{P}(n,p)$, which is the reason why we need the additional $\log n$ factor.

Finally, we remark that by weighting the elements $a \in \mathcal{P}(n)$ according to the number of $[j,r]$-chains meeting a, the above proof may easily be modified to yield LYM versions of Theorems 1 and 2. The following result is the analogue of Theorem 1.

Theorem 9 *Let $\varepsilon > 0$ be given. Suppose that $pn/\log n \to \infty$ and that $A \subseteq \mathcal{P}(n,p)$ is an antichain with $\varepsilon n \le |a| \le (1-\varepsilon)n$ for all $a \in A$. Then* a.s.

$$\sum_{a\in A} \frac{1}{\binom{n}{|a|}} \le (1+o(1))\,p.$$

Note that it is necessary to impose some condition on $|a|$, since for instance if p is not much larger than the above threshold, the proof of Proposition 4 shows that *a.s.* one may pick an antichain consisting of all elements in the first level and almost all of the elements in the second level of $\mathcal{P}(n,p)$, and whose weight is $(2+o(1))p$.

4 Acknowledgements

I would like to thank Yoshi Kohayakawa and the referee for a careful reading of the proof.

References

[1] N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley, 1992.

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

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

[4] Y. Kohayakawa, B. Kreuter, and D. Osthus, The length of random subsets of Boolean lattices, *Random Struct. Alg.*, to appear.

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

[6] D. Lubell, A short proof of Sperner's Lemma, *J. Combin. Theory* **1**, 299 (1966).

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