Supersaturation, counting, and randomness in forbidden subposet problems

Dániel Gerbner[1]   Dániel T. Nagy[1]   Balázs Patkós[1,2]   Máté Vizer[1]

[1] Alfréd Rényi Institute of Mathematics, Budapest
[2] Moscow Institute of Physics and Technology

Abstract

In the area of forbidden subposet problems we look for the largest possible size $La(n, P)$ of a family $\mathcal{F} \subseteq 2^{[n]}$ that does not contain a forbidden inclusion pattern described by P. The main conjecture of the area states that for any finite poset P there exists an integer $e(P)$ such that $La(n, P) = (e(P) + o(1))\binom{n}{\lfloor n/2 \rfloor}$.

In this paper, we formulate three strengthenings of this conjecture and prove them for some specific classes of posets. (The parameters $x(P)$ and $d(P)$ are defined in the paper.)

- For any finite connected poset P and $\varepsilon > 0$, there exists $\delta > 0$ and an integer $x(P)$ such that for any n large enough, and $\mathcal{F} \subseteq 2^{[n]}$ of size $(e(P) + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$, $\mathcal{F}$ contains at least $\delta n^{x(P)}\binom{n}{\lfloor n/2 \rfloor}$ copies of P.

- The number of P-free families in $2^{[n]}$ is $2^{(e(P)+o(1))\binom{n}{\lfloor n/2 \rfloor}}$.

- For any finite poset P, there exists a positive rational $d(P)$ such that if $p = \omega(n^{-d(P)})$, then the size of the largest P-free family in $\mathcal{P}(n, p)$ is $(e(P) + o(1))p\binom{n}{\lfloor n/2 \rfloor}$ with high probability.

1 Introduction

Extremal set theory starts with the seminal result of Sperner [21] that was generalized by Erdős [7] as follows: if a family $\mathcal{F} \subseteq 2^{[n]}$ of sets does not contain a nested sequence $F_1 \subsetneq F_2 \subsetneq \cdots \subsetneq F_{k+1}$ (such nested sequences are called *chains of length $k + 1$* or *$(k + 1)$-chains* for short), then its size cannot exceed that of the union of k middle levels of $2^{[n]}$, i.e., $|\mathcal{F}| \leq \sum_{i=1}^{k} \binom{n}{\lfloor \frac{n-k}{2} \rfloor + i}$. This theorem has many applications and several of its variants have been investigated.

In the early 80's, Katona and Tarján [13] introduced the following general framework to study set families avoiding some fixed inclusion patterns: we say that a subfamily $\mathcal{G}$ of $\mathcal{F}$ is a *(non-induced) copy* of a poset $(P, \leq)$ in $\mathcal{F}$, if there exists a bijection $i : P \to \mathcal{G}$ such that if $p, q \in P$

with $p \leq q$, then $i(p) \subseteq i(q)$. If i satisfies the property that for $p, q \in P$ we have $p \leq q$ if and only if $i(p) \subseteq i(q)$, then $\mathcal{G}$ is called an *induced copy* of P in $\mathcal{F}$. If $\mathcal{F}$ does not contain any (induced) copy of P, the $\mathcal{F}$ is said to be *(induced) P-free*. The largest possible size of a(n induced) P-free family $\mathcal{F} \subseteq 2^{[n]}$ is denoted by $La(n, P)$ $(La^*(n, P))$. Let P_k denote the k-chain, then the result of Erdős mentioned above determines $La(n, P_{k+1})$. These parameters have attracted the attention of many researchers, and there are widely believed conjectures in the area (see Conjecture 1) that appeared first in [3] and [11], giving the asymptotics of $La(n, P)$ and $La^*(n, P)$.

Let $e(P)$ denote the maximum integer m such that for any $i \leq n$, the family $\binom{[n]}{i+1} \cup \binom{[n]}{i+2} \cup \cdots \cup \binom{[n]}{i+m}$ is P-free. Similarly, let $e^*(P)$ denote the maximum integer m such that for any $i \leq n$, the family $\binom{[n]}{i+1} \cup \binom{[n]}{i+2} \cup \cdots \cup \binom{[n]}{i+m}$ is induced P-free.

Conjecture 1.
(i) $La(n, P) = (e(P) + o(1))\binom{n}{\lfloor n/2 \rfloor}$.
(ii) $La^*(n, P) = (e^*(P) + o(1))\binom{n}{\lfloor n/2 \rfloor}$.

Conjecture 1 has been verified for several classes of posets, but is still open in general. For more results on the $La(n, P)$ function, see Chapter 7 of [8], and see other chapters for more background on the generalizations considered in this paper.

After determining (the asymptotics of) the extremal size and the structure of the extremal families, one may continue in several directions. *Stability* results state that all P-free families having almost extremal size must be very similar in structure to the middle $e(P)$ levels of $2^{[n]}$. *Supersaturation* problems ask for the minimum number of copies of P that a family $\mathcal{F} \subseteq 2^{[n]}$ of size $La(n, P) + E$ may contain. This is clearly at least E, but usually one can say much more. *Counting* problems ask to determine the number of P-free families in $2^{[n]}$. As any subfamily of a P-free family is P-free, therefore the number of P-free families is at least $2^{La(n,P)}$. The question is how many more such families there are. Finally, one can address *random versions* of the forbidden subposet problem. Let $\mathcal{P}(n, p)$ denote the probability space of all subfamilies of $2^{[n]}$ such that for any $F \subseteq [n]$, the probability that F belongs to $\mathcal{P}(n, p)$ is p, independently of any other set F'. What is the size of the largest P-free subfamily of $\mathcal{P}(n, p)$ with high probability[1]? Clearly, for $p = 1$, this is $La(n, P)$. For other values of p, an obvious construction is to take a P-free subfamily of $2^{[n]}$, and then the sets that are in $\mathcal{P}(n, p)$ form a P-free family. Taking the $e(P)$ middle levels shows that the size of the largest P-free family in $\mathcal{P}(n, P)$ is at least $p(e(P) + o(1))\binom{n}{\lfloor n/2 \rfloor}$ w.h.p.. For what values of p does this formula give the asymptotically correct answer?

In this paper, we will consider supersaturation, counting and random versions of the forbidden subposet problem, mostly focusing on supersaturation results. We will propose three

---

[1] we say that a sequence of events $E_1, E_2, \ldots, E_n, \ldots$ holds with high probability (or w.h.p., in short) if $\mathbb{P}(E_n)$ tends to 1 as n tends to infinity

strengthenings of Conjecture 1 and prove them for some classes of posets. In the remainder of the introduction, we state our results and also what was known before.

The supersaturation version of Sperner's problem is to determine the minimum number of pairs $F \subsetneq F'$ over all subfamilies of $2^{[n]}$ of given size. We say that a family $\mathcal{F}$ is *centered* if it consists of the sets closest to $n/2$. More precisely, if $F \in \mathcal{F}$ and $||G| - n/2| < ||F| - n/2|$ imply $G \in \mathcal{F}$. Kleitman [14] proved that among families of cardinality m, centered ones contain the smallest number of copies of P_2. He conjectured that the same holds for any P_k. After several partial results, e.g. [2, 5, 6], the conjecture was confirmed by Samotij [20]. The following is a consequence of the result of Samotij. We will only use it with $k = 2$, i.e. the result of Kleitman.

Theorem 2. *For any k, t with $k - 1 \le t$ and $\varepsilon > 0$ there exists $n_{k,t,\varepsilon}$ such that if $n \ge n_{k,t,\varepsilon}$, then any family $\mathcal{F} \subseteq 2^{[n]}$ of size at least $(t + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\varepsilon \frac{n^t}{2^{t+1}}\binom{n}{\lfloor n/2 \rfloor}$ chains of length k.*

Apart from the above, the only supersaturation result in the area of forbidden subposet problems is due to Patkós [19]. It gives the minimum number of copies of the butterfly poset[2] B in families of size $La(n, B) + E$ for small values of E.

We will investigate the number of copies of P created when the number of additional sets compared to a largest P-free family is proportional to the size of the middle level $\binom{[n]}{\lfloor n/2 \rfloor}$. Let $M(n, P)$ denote the number of copies of P in the $e(P) + 1$ middle levels of $2^{[n]}$, and let $M^*(n, P)$ denote the number of induced copies of P in the $e^*(P) + 1$ middle levels of $2^{[n]}$. The *Hasse diagram* of a poset P is the directed graph with vertex set P and for $p, q \in P$, (pq) is an arc in the Hasse diagram if $p < q$ and there does not exist $z \in P$ with $p < z < q$. We say that P is *connected*, if its Hasse diagram (as a digraph) is weakly connected, i.e., we cannot partition its vertices into two sets such that there is no arc between those sets. The undirected Hasse diagram is the undirected graph obtained from the Hasse diagram by removing orientations of all arcs.

Proposition 3. *For any connected poset P on at least two elements there exist positive integers $x(P)$ and $x^*(P)$ such that $M(n, P) = \Theta\left(n^{x(P)}\binom{n}{\lfloor n/2 \rfloor}\right)$ and $M^*(n, P) = \Theta\left(n^{x^*(P)}\binom{n}{\lfloor n/2 \rfloor}\right)$ hold.*

Proof. The proofs of the two statements are analogous, so we include only that of the non-induced version. For a copy $\mathcal{G}$ of P with all sets from the $e(P) + 1$ middle levels of $2^{[n]}$, let $A_{\mathcal{G}} = \cap_{G \in \mathcal{G}} G$, $B_{\mathcal{G}} = \cup_{G \in \mathcal{G}} G$ and $m_{\mathcal{G}} = |A_{\mathcal{G}}|$, $M_{\mathcal{G}} = |B_{\mathcal{G}}|$. Let us define $x(P)$ now. Let $x(P) := \max_{\mathcal{G}}\{M_{\mathcal{G}} - m_{\mathcal{G}}\}$, where the maximum runs through all the copies $\mathcal{G}$ of P with all sets from the $e(P) + 1$ middle levels of $2^{[n]}$.

We claim that for any such $\mathcal{G}$, we have $M_{\mathcal{G}} - m_{\mathcal{G}} \le e(P)|P|$ (in other words $x(P) \le e(P)|P|$). Indeed, as P is connected, we can go through its elements in an order such that every element is in relation with at least one of the earlier elements. As $\mathcal{G}$ is from the $e(P) + 1$ middle levels, this means that the new element is a set that is either contained in, or contains an earlier set, thus it

---

[2] the poset on four elements $a, b < c, d$

3

differs from that set in at most $e(P)$ elements. In the first case, this new element decreases $m_{\mathcal{G}}$ by at most $e(P)$, in the second case it increases $M_{\mathcal{G}}$ by at most $e(P)$, so we are done.

Similarly, one can show $|m_{\mathcal{G}} - n/2|, |M_{\mathcal{G}} - n/2| \leq e(P)|P|$. Clearly, for any $A \subseteq B$ with $|B| - |A| \leq e(P)|P|$ there is at most a fixed constant number of copies $\mathcal{G}$ of P such that $A = A_{\mathcal{G}}$ and $B = B_{\mathcal{G}}$. Finally, the number of pairs $A \subseteq B$, with $|B \setminus A| \leq x(P)$, $||A| - n/2| \leq e(P)|P|$ and $||B| - n/2| \leq e(P)|P|$ is at most $C\binom{n}{\lfloor n/2 \rfloor}\binom{\lfloor n/2 \rfloor + e(P)|P|}{x(P)}$. This yields $M(n, P) = O(n^{x(P)}\binom{n}{\lfloor n/2 \rfloor})$.

For the lower bound, fix a copy $\mathcal{G}$ with $|B_{\mathcal{G}} \setminus A_{\mathcal{G}}| = x(P)$. Clearly, for any $A' \subseteq B'$ with $|A'| = |A_{\mathcal{G}}|, |B'| = |B_{\mathcal{G}}|$ there exists a permutation π of $[n]$ with $\pi(A_{\mathcal{G}}) = A'$ and $\pi(B_{\mathcal{G}}) = B'$. Therefore, such permutations π map $\mathcal{G}$ into distinct copies of P. Their number is clearly at least $\binom{n}{|A_{\mathcal{G}}|}\binom{n - |A_{\mathcal{G}}|}{|B_{\mathcal{G}} \setminus A_{\mathcal{G}}|} \geq c \cdot n^{x(P)}\binom{n}{\lfloor n/2 \rfloor}$ for some positive constant c. $\square$

Now we can state the first generalization of Conjecture 1.

Conjecture 4.

(i) For every poset P and $\varepsilon > 0$ there exists $\delta > 0$ such that if $\mathcal{F} \subseteq 2^{[n]}$ is of size at least $(e(P) + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$, then $\mathcal{F}$ contains at least $\delta \cdot M(n, P)$ many copies of P.

(ii) For every poset P and $\varepsilon > 0$ there exists $\delta > 0$ such that if $\mathcal{F} \subseteq 2^{[n]}$ is of size at least $(e^*(P) + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$, then $\mathcal{F}$ contains at least $\delta \cdot M^*(n, P)$ many induced copies of P.

We will prove Conjecture 4 for several classes of tree posets. A poset T is a *tree poset*, if its undirected Hasse diagram is a tree. Note that for any tree poset T of height 2, we have $x(T) = x^*(T) = |T| - 1$.

Theorem 5. *Let T be any height 2 tree poset of $t + 1$ elements. Then for any $\varepsilon > 0$ there exist $\delta > 0$ and n_0 such that for any $n \geq n_0$ any family $\mathcal{F} \subseteq 2^{[n]}$ of size $|\mathcal{F}| \geq (1 + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta n^t \binom{n}{\lfloor n/2 \rfloor}$ copies of T.*

We say that a tree poset T is *upward (downward) monotone*, if for any $x \in T$ there exists at most 1 element $y \in T$ with $y \prec x$ ($x \prec y$). A tree poset is called *monotone*, if it is either upward or downward monotone.

Theorem 6. *For any monotone tree poset T and $\varepsilon > 0$, there exist $\delta > 0$ and n_0 such that for any $n \geq n_0$ any family $\mathcal{F} \subseteq 2^{[n]}$ of size $|\mathcal{F}| \geq (h(T) - 1 + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta n^{x(T)}\binom{n}{\lfloor n/2 \rfloor}$ copies of T.*

The complete multipartite poset $K_{r_1, r_2, \ldots, r_\ell}$ is a poset on $\sum_{i=1}^{\ell} r_i$ elements $a_{i,j}$ with $i = 1, 2, \ldots, \ell$, $j = 1, 2, \ldots, r_i$ such that $a_{i,j} < a_{i',j'}$ if and only if $i < i'$. The poset $K_{1,r}$ is usually denoted by $\vee_r$, and the poset $K_{r,1}$ is denoted by $\wedge_r$. The poset $K_{s,1,t}$ is a tree poset with $x(K_{s,1,t}) = x^*(K_{s,1,t}) = s + t$.

Theorem 7. *For any $s, t \in \mathbb{N}$ and $\varepsilon > 0$ there exist $n_0 = n_{\varepsilon,s,t}$ and $\delta > 0$ such that any $\mathcal{F} \subseteq 2^{[n]}$ of size at least $(2 + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ with $n \geq n_0$ contains at least $\delta n^{s+t}\binom{n}{\lfloor n/2 \rfloor}$ induced copies of $K_{s,1,t}$.*

We will consider the supersaturation problem for the generalized diamond D_s, i.e., the poset on $s+2$ elements with $a < b_1, b_2, \ldots, b_s < c$. For any integer $s \geq 2$, let us define $m_s = \lceil \log_2(s+2) \rceil$ and $m_s^* = \min\{m : s \leq \binom{m}{\lceil m/2 \rceil}\}$. Clearly, for any integer $s \geq 2$, we have $e(D_s) = x(D_s) = m_s$ and $e^*(D_s) = x^*(D_s) = m_s^*$. The next theorem establishes a lower bound that is less by a factor of $\sqrt{n}$ than what Conjecture 4 states for diamond posets D_s for infinitely many s.

Theorem 8.

(i) If $s \in [2^{m_s-1} - 1, 2^{m_s} - \binom{m_s}{\lceil \frac{m_s}{2} \rceil} - 1]$, then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that every $\mathcal{F} \subseteq 2^{[n]}$ with $|\mathcal{F}| \geq (m_s + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta \cdot n^{m_s - 0.5}\binom{n}{\lfloor n/2 \rfloor}$ copies of D_s.

(ii) For any $\varepsilon > 0$ there exists a $\delta > 0$ such that every $\mathcal{F} \subseteq 2^{[n]}$ with $|\mathcal{F}| \geq (4 + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta \cdot n^{3.5}\binom{n}{\lfloor n/2 \rfloor}$ induced copies of D_4.

(iii) For any constant c with $1/2 < c < 1$ there exists an integer s_c such that if $s \geq s_c$ and $s \leq c\binom{m_s^*}{\lfloor m_s^*/2 \rfloor}$, then the following holds: for any $\varepsilon > 0$ there exists a $\delta > 0$ such that every $\mathcal{F} \subseteq 2^{[n]}$ with $|\mathcal{F}| \geq (m_s^* + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta \cdot n^{m_s^* - 0.5}\binom{n}{\lfloor n/2 \rfloor}$ induced copies of D_s.

Let us turn our attention to counting (induced) P-free families. As we mentioned earlier, every subfamily of a P-free family is P-free, therefore $2^{La(n,P)} \geq 2^{(e(P)+o(1))\binom{n}{\lfloor n/2 \rfloor}}$ is a lower bound on the number of such families. Determining the number of P_2-free families has attracted a lot of attention. The upper bound $2^{(1+o(1))\binom{n}{\lfloor n/2 \rfloor}}$, asymptotically matching in the exponent the trivial lower bound was obtained by Kleitman [15]. After several improvements, Korshunov [16] determined asymptotically the number of P_2-free families.

Conjecture 9. (i) The number of P-free families in $2^{[n]}$ is $2^{(e(P)+o(1))\binom{n}{\lfloor n/2 \rfloor}}$.

(ii) The number of induced P-free families in $2^{[n]}$ is $2^{(e^*(P)+o(1))\binom{n}{\lfloor n/2 \rfloor}}$.

Theorem 10. (i) The number of induced $\vee_r$-free families is $2^{(1+o(1))\binom{n}{\lfloor n/2 \rfloor}}$.

(ii) The number of induced $K_{s,1,t}$-free families in $2^{[n]}$ is $2^{(2+o(1))\binom{n}{\lfloor n/2 \rfloor}}$.

As every height 2 poset P is a non-induced subposet of $K_{|P|,1,|P|}$, Conjecture 9 (i) is an immediate consequence of Theorem 10 for those height 2 posets P for which $e(P) = 2$.

Finally, we turn to random versions of forbidden subposet problems. The probabilistic version of Sperner's theorem was proved by Balogh, Mycroft, and Treglown [1] and Collares and Morris [17], independently. It states that if $p = \omega(1/n)$, then the largest antichain in $\mathcal{P}(n,p)$ is of size $(1 + o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p.. This is sharp in the sense that if $p = o(1/n)$ then the asymptotics is different. Note that as any k-Sperner family is the union of k antichains, the analogous statement holds for k-Sperner families in $\mathcal{P}(n,p)$. Both papers used the container method. Hogenson in her PhD thesis [12] adapted the method of Balogh, Mycroft, and Treglown to obtain the same results for non-induced $\vee_r$-free families.

Let us state a general proposition that gives a range of p when one can have a P-free family in $\mathcal{P}(n,p)$ that is larger than $p(e(P) + o(1))\binom{[n}{\lfloor n/2 \rfloor}$.

Proposition 11. *For any finite connected poset P, the following statements hold.*

(i) If $p = o(n^{-\frac{x(P)}{|P|-1}})$, then the largest P-free family in $\mathcal{P}(n,p)$ has size at least $(e(P)+1-o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p..

(ii) If $p = o(n^{-\frac{x^*(P)}{|P|-1}})$, then the largest induced P-free family in $\mathcal{P}(n,p)$ has size at least $(e^*(P)+1-o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p..

Proof. We only prove (i), the proof of (ii) is similar. Let us denote the random family of the $e(P)+1$ middle levels after keeping any of its sets with probability p by $\mathcal{M}_p = \mathcal{M} \cap \mathcal{P}(n,p)$. Clearly, $\mathbb{E}(|\mathcal{M}_p|) = (e(P)+1+o(1))p\binom{n}{\lfloor n/2 \rfloor}$ and thus we have $|\mathcal{M}_p| = (e(P)+1+o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p.. Let X be the random variable that denotes the number of copies of P in $\mathcal{M}_p$. Then we have $\mathbb{E}(X) = \Theta(p^{|P|}n^{x(P)}\binom{n}{\lfloor n/2 \rfloor})$. By the assumption on p, we have $p^{|P|}n^{x(P)}\binom{n}{\lfloor n/2 \rfloor} = o(p\binom{n}{\lfloor n/2 \rfloor})$, and so $X = o(|\mathcal{M}_p|)$ w.h.p., and thus by removing the copies of P from $\mathcal{M}_p$, we obtain a P-free family in $\mathcal{P}(n,p)$ of size $(e(P)+1-o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p.. $\square$

If $\mathcal{M}_p$ does not contain a subposet P' of P, then it is P-free, thus we have the following.

Corollary 12. *For any finite poset P, let $d(P) = \min \frac{x(P')}{|P'|-1}$, where P' runs through all connected subposets P' of P with $e(P) = e(P')$. Similarly, let $d^*(P) = \min \frac{x^*(P')}{|P'|-1}$, where P' runs through all connected subposets P' of P with $e^*(P) = e^*(P')$. Then the following statements hold.*

(i) If $p = o(n^{-d(P)})$, then the largest P-free family in $\mathcal{P}(n,p)$ has size at least $(e(P)+1-o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p.

(ii) If $p = o(n^{-d^*(P)})$, then the largest induced P-free family in $\mathcal{P}(n,p)$ has size $(e^*(P)+1-o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p.

We conjecture that the bounds above are sharp.

Conjecture 13. *For any finite connected poset P the following statements hold.*

(i) If $p = \omega(n^{-d(P)})$, then the largest P-free family in $\mathcal{P}(n,p)$ has size $(e(P)+o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p..

(ii) If $p = \omega(n^{-d^*(P)})$, then the largest induced P-free family in $\mathcal{P}(n,p)$ has size $(e^*(P)+o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p..

Theorem 14. *If $p = \omega(1/n)$, then the following are true.*

(i) For any integer $r \geq 0$, the largest induced $\vee_{r+1}$-free family in $\mathcal{P}(n,p)$ has size $(1+o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p..

(ii) For any pair $s,t \geq 1$ of integers, the largest induced $K_{s,1,t}$-free family in $\mathcal{P}(n,p)$ has size $(2+o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p..

The structure of the paper is as follows: in the next section, we gather some earlier results that will be used as tools in the proofs of our theorems. Also, we will obtain an induced $\vee_{r+1}$-free container lemma. Section 3, in three subsections, contain the proofs of our theorems.

2   Preliminaries

Here and in the next section, we will assume that n is large enough whenever it is necessary.

Lemma 15. *For any $1 \le l \le n/2$ we have $\sum_{i=0}^{l-1} \binom{n}{i} \le 2\sqrt{n}\binom{n}{l}$.*

Proof. Let $m = \lfloor \sqrt{n} \rfloor$ and observe that for any $k < n/2$ we have

$$\frac{\binom{n}{k-m}}{\binom{n}{k}} \le \frac{\binom{n}{\lfloor n/2 \rfloor - m}}{\binom{n}{\lfloor n/2 \rfloor}} \le \prod_{i=1}^{m} \frac{\lfloor n/2 \rfloor - i + 1}{\lceil n/2 \rceil + i} \le e^{-\sum_{i=1}^{m} \frac{2i}{n}} \le e^{-1}.$$

So dividing $\sum_{i=0}^{l-1} \binom{n}{i}$ into m subsums depending on the residue of $i \bmod m$, we obtain subsums that can be upper bounded by geometric progressions of quotient e^{-1}. $\qquad\square$

We denote by $\mathbf{C}_k$ the set of chains of length $k+1$, i.e. the set of maximal chains in $2^{[k]}$.

Lemma 16 (Griggs, Li, Lu, in the proof of Theorem 2.5 in [10])**.**
   If $s \in [2^{m_s-1} - 1, 2^{m_s} - \binom{m_s}{\lceil \frac{m_s}{2} \rceil} - 1]$, and $\mathcal{G} \subseteq 2^{[k]}$ is a D_s-free family of sets, then the number of pairs $(G, \mathcal{C})$ with $G \in \mathcal{G} \cap \mathcal{C}$ and $\mathcal{C} \in \mathbf{C}_k$ is at most $m_s k!$.

Lemma 17 (Patkós [18])**.**
   (i) Let $\mathcal{G} \subseteq 2^{[k]}$ be a family of sets such that any antichain $\mathcal{A} \subset \mathcal{G}$ has size at most 3. Then the number of pairs $(G, \mathcal{C})$ with $G \in \mathcal{G} \cap \mathcal{C}$ and $\mathcal{C} \in \mathbf{C}_k$ is at most $4k!$.
   (ii) For any constant c with $1/2 < c < 1$ there exists an integer s_c such that if $s \ge s_c$ and $s \le c\binom{m_s^*}{\lceil m_s^*/2 \rceil}$, then the following holds: if $\mathcal{G} \subseteq 2^{[k]}$ is a family of sets such that any antichain $\mathcal{A} \subset \mathcal{G}$ has size less than s, then the number of pairs $(G, \mathcal{C})$ with $G \in \mathcal{G} \cap \mathcal{C}$ and $\mathcal{C} \in \mathbf{C}_k$ is at most $m_s^* k!$.

The special case $r = 1$ of the next theorem was proved by Carroll and Katona [4].

Theorem 18 (Griggs, Li [9])**.**
   If $\mathcal{F} \subseteq 2^{[n]}$ is an induced $\vee_{r+1}$-free family, then $|\mathcal{F}| \le (1 + \frac{2r}{n} + O(\frac{r}{n^2}))\binom{n}{\lfloor n/2 \rfloor}$ holds.

We remark that we will use the above theorem with $r = \Theta(n)$.

Theorem 19.
   For every $\varepsilon > 0$ and $t, r \in \mathbb{Z}^+$, if n is large enough and $\mathcal{F} \subseteq 2^{[n]}$ is of size at least $(t+\varepsilon)\binom{n}{\lfloor n/2 \rfloor}$, then there exists $F \in \mathcal{F}$ such that F is the bottom element of an induced copy of $\vee_{\delta n^t}$ with $\delta = \frac{\varepsilon}{2^{t+3}t!}$.

Proof. As $|\{G \subseteq [n] : ||G| - n/2| \geq n^{2/3}\}| \leq \frac{1}{n^2}\binom{n}{\lfloor n/2 \rfloor}$, we can assume that $n/2 - n^{2/3} \leq |F| \leq n + n^{2/3}$ holds for every $F \in \mathcal{F}$ and $|\mathcal{F}| \geq (t + \varepsilon/2)\binom{n}{\lfloor n/2 \rfloor}$. The Lubell-function $\lambda_n(\mathcal{F}) := \sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}}$ of a family $\mathcal{F} \subseteq 2^{[n]}$ of sets is the average number of sets in $\mathcal{F} \cap \mathcal{C}$ over all maximal chains $\mathcal{C} \in \mathbf{C}_n$. Clearly, $|\mathcal{F}| \geq (t + \varepsilon/2)\binom{n}{\lfloor n/2 \rfloor}$ implies $\lambda_n(\mathcal{F}) \geq t + \varepsilon/2$.

Let $\mathbf{C}_F$ be the collection of maximal chains $\mathcal{C} \in \mathbf{C}_n$ such that F is the smallest set in $\mathcal{F} \cap \mathcal{C}$, and let $\mathbf{C}_\emptyset$ be the collection of those maximal chains that avoid $\mathcal{F}$. Let us partition $\mathbf{C}_n$ into $\mathbf{C}_\emptyset \cup \bigcup_{F \in \mathcal{F}} \mathbf{C}_F$. Writing $\mathcal{F}_F = \{F' \setminus F : F' \in \mathcal{F}, F \subseteq F'\}$, we obtain $\lambda_n(\mathcal{F}) = \sum_{F \in \mathcal{F}} \frac{|\mathbf{C}_F|}{n!}\lambda_{n-|F|}(\mathcal{F}_F)$. This means that for some $F \in \mathcal{F}$, we must have $\lambda_{n-|F|}(\mathcal{F}_F) \geq t + \varepsilon/2$. If $\mathcal{F}$ does not contain any induced copy of $\vee_{\delta n^t}$, then for any $i \geq t$ we have $|\{F' \in \mathcal{F}_F : |F''| = i\}| < \delta n^t$, thus writing $\mathcal{F}_{F,1} = \{G \in \mathcal{F}_F : |G| \leq t - 1\}$ and $\mathcal{F}_{F,2} = \mathcal{F}_F \setminus \mathcal{F}_{F,1}$ we obtain

$$t + \varepsilon/2 \leq \lambda_{n-|F|}(\mathcal{F}_F) = \lambda_{n-|F|}(\mathcal{F}_{F,1}) + \lambda_{n-|F|}(\mathcal{F}_{F,2}) \leq$$

$$\leq t + \delta n^t \sum_{i=t}^{2n^{2/3}} \frac{1}{\binom{n-|F|}{i}} < t + \frac{2\delta n^t}{\binom{n/2-n^{2/3}}{t}} \leq t + 2^{t+2}\delta t!.$$

This is a contradiction if $\delta \leq \frac{\varepsilon}{2^{t+3}t!}$. $\qquad\square$

Hogenson [12], altering a proof of Balogh, Mycroft, and Treglown [1], obtained a container lemma for non-induced $\vee_{r+1}$-free families. With the help of Theorem 19, we can validate it for induced $\vee_{r+1}$-free families.

For a family $\mathcal{G}$ of sets in $2^{[n]}$ and a set $F \subseteq 2^{[n]}$, we introduce $U_\mathcal{G}(F) = \{G \in \mathcal{G} : F \subseteq G\}$. Previous proofs used the $\mathcal{G}$-*degree* of F, the size of $U_\mathcal{G}(F)$. Observe that $\mathcal{F}$ is $\vee_{r+1}$-free if and only if the $\mathcal{F}$-degree of every $F \in \mathcal{F}$ is at most r. We introduce the $\mathcal{G}$-*weight* of F as the size of the largest antichain in $U_\mathcal{G}(F)$. Observe that $\mathcal{F}$ is induced $\vee_{r+1}$-free if and only if the $\mathcal{F}$-weight of every $F \in \mathcal{F}$ is at most r. Now we state the appropriate container lemma.

Theorem 20. *Let $t, r \in \mathbb{Z}^+$ and $\varepsilon \leq \frac{1}{(2t)^{t+1}}$ and assume n is large enough. Then there exist functions $f : \binom{2^{[n]}}{\leq (r+1)2^n n^{-(t+0.9)}} \to \binom{2^{[n]}}{(t+1+\varepsilon)\binom{n}{\lfloor n/2 \rfloor}}$ and $g : \binom{2^{[n]}}{\leq (r+1)\frac{t+2}{\varepsilon^2 n^t}\binom{n}{\lfloor n/2 \rfloor}} \to \binom{2^{[n]}}{(t+\varepsilon)\binom{n}{\lfloor n/2 \rfloor}}$ such that for any induced $\vee_{r+1}$-free family $\mathcal{F} \subseteq 2^{[n]}$ there exist disjoint subfamilies $\mathcal{H}_1, \mathcal{H}_2 \subseteq \mathcal{F}$ such that $(\mathcal{H}_1 \cup \mathcal{H}_2) \cap g(\mathcal{H}_1 \cup \mathcal{H}_2) = \emptyset$, $\mathcal{H}_2 \subseteq f(\mathcal{H}_1)$ and $\mathcal{F} \subseteq \mathcal{H}_1 \cup \mathcal{H}_2 \cup g(\mathcal{H}_1 \cup \mathcal{H}_2)$.*

Proof. The proof uses the standard graph container algorithm with some modifications. First, we fix an ordering $S_1, S_2, \ldots, S_{2^n}$ of $2^{[n]}$ and also an ordering $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{2^{2^n}}$ of $2^{2^{[n]}}$. The input of the algorithm is an induced $\vee_{r+1}$-free family $\mathcal{F} \subseteq 2^{[n]}$, and in two phases it outputs $\mathcal{H}_1$, $\mathcal{H}_2$, $f(\mathcal{H}_1)$ and $g(\mathcal{H}_1 \cup \mathcal{H}_2)$ as follows.

At the beginning we set $\mathcal{G}^0 = 2^{[n]}$, $\mathcal{H}_1^0 = \mathcal{H}_2^0 = \emptyset$ and we start at Phase I.

Later in the ith round (for $i = 1, 2, \ldots$) we always pick a $G_i \in \mathcal{G}^{i-1}$ with largest $\mathcal{G}^{i-1}$-weight. If there are several sets G with the same (largest) $\mathcal{G}^{i-1}$-weight, we pick the one appearing first in our fixed ordering of $2^{[n]}$.

8

Phase I.

- If $G_i \notin \mathcal{F}$, we set $\mathcal{G}^i = \mathcal{G}^{i-1} \setminus \{G\}$ and $\mathcal{H}_1^i = \mathcal{H}_1^{i-1}$, $\mathcal{H}_2^i = \mathcal{H}_2^{i-1}$.

- If $G_i \in \mathcal{F}$, and the $\mathcal{G}^{i-1}$-weight of G_i is more than $n^{t+0.9}$, then we pick the largest antichain $\mathcal{A}^i$ in $U_{\mathcal{G}^{i-1}}(G_i)$. If there are multiple such antichains, we pick the one with the smallest index in our fixed ordering of $2^{2^{[n]}}$ and set $\mathcal{G}^i = \mathcal{G}^{i-1} \setminus (\mathcal{A}^i \cup \{G_i\})$, $\mathcal{H}_1^i = \mathcal{H}_1^{i-1} \cup [(\mathcal{A}^i \cap \mathcal{F}) \cup \{G_i\}]$ and $\mathcal{H}_2^i = \mathcal{H}_2^{i-1}$.

- If $G_i \in \mathcal{F}$, and the $\mathcal{G}^{i-1}$-weight of G_i is at most $n^{t+0.9}$, then Phase I is ended, we keep $\mathcal{G}^i = \mathcal{G}^{i-1}$, $\mathcal{H}_2^i = \mathcal{H}_2^{i-1}$, set $\mathcal{H}_1 = \mathcal{H}_1^{i-1}$ (from here on, $\mathcal{H}_1^i$ does not change) and define $f(\mathcal{H}_1) = \mathcal{G}^i$. And jump to Phase II.

Phase II.

- If $G_i \notin \mathcal{F}$, we set $\mathcal{G}^i = \mathcal{G}^{i-1} \setminus \{G\}$ and $\mathcal{H}_1^i = \mathcal{H}_1^{i-1}$, $\mathcal{H}_2^i = \mathcal{H}_2^{i-1}$.

- If $G_i \in \mathcal{F}$, and the $\mathcal{G}^{i-1}$-weight of G_i is more than $\varepsilon^2 n^t$, then we again take the largest antichain $\mathcal{A}^i$ in $U_{\mathcal{G}^{i-1}}(G_i)$ with the smallest index in our fixed ordering of $2^{2^{[n]}}$ and set $\mathcal{G}^i = \mathcal{G}^{i-1} \setminus (\mathcal{A}^i \cup \{G_i\})$, $\mathcal{H}_2^i = \mathcal{H}_2^{i-1} \cup [(\mathcal{A}^i \cap \mathcal{F}) \cup \{G_i\}]$.

- If $G_i \in \mathcal{F}$, and the $\mathcal{G}^{i-1}$-weight of G_i is at most $\varepsilon^2 n^t$, then Phase II and the algorithm is ended, set $\mathcal{H}_2 = \mathcal{H}_2^{i-1}$ and define $g(\mathcal{H}_1 \cup \mathcal{H}_2) = \mathcal{G}^{i-1}$.

Observe the following:

- whenever we include sets in $\mathcal{H}_1^i$, then the number of such sets is at most $r + 1$ (as $\mathcal{F}$ is $\vee_{r+1}$-free), and the number of sets removed from $\mathcal{G}^{i-1}$ is at least $n^{t+0.9}$, so

$$|\mathcal{H}_1| \leq (r+1)2^n / n^{t+0.9};$$

- at the end of Phase I, $\mathcal{G}^i$ does not contain any induced copies of $\vee_{n^{t+0.9}}$, so, by Theorem 19,

$$|f(\mathcal{H}_1)| = |\mathcal{G}^i| \leq (t + 1 + \varepsilon)\binom{n}{\lfloor n/2 \rfloor};$$

- the above two bullet points and the threshold for Phase II imply that

$$|\mathcal{H}_2| \leq \frac{r+1}{\varepsilon^2 n^t}|f(\mathcal{H}_1)| \leq (r+1)\frac{t+2}{\varepsilon^2 n^t}\binom{n}{\lfloor n/2 \rfloor};$$

- Theorem 19 implies that at the end of Phase II,

$$|g(\mathcal{H}_1 \cup \mathcal{H}_2)| = |\mathcal{G}^i| \leq (t + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}.$$

All what remains to prove is that the function f and g are well defined, i.e., if for two distinct $\vee_{r+1}$-free families $\mathcal{F}$ and $\mathcal{F}'$ the algorithm outputs the same $\mathcal{H}_1$, then $f(\mathcal{H}_1)$ is defined the same, and if in addition $\mathcal{H}_1 \cup \mathcal{H}_2$ is the same for both run of the algorithm, then so is $g(\mathcal{H}_1 \cup \mathcal{H}_2)$. We claim more: the families $\mathcal{G}^i, \mathcal{H}_1^i, \mathcal{H}_2^i$ are the same for both runs for all values of $i = 0, 1, \ldots$. This is certainly true for $i = 0$. Then by induction, if this holds for some i, then G^i is the same for both run. Therefore, due to the fixed ordering of $2^{[n]}$, the algorithm considers the same set $G_{i+1} \in \mathcal{G}^i$ in step $i + 1$. As G_{i+1} will be removed from $\mathcal{G}^i$ in all cases, therefore it cannot happen that G_{i+1} belongs to exactly one of $\mathcal{F}, \mathcal{F}'$. If $G_{i+1} \notin \mathcal{F}, \mathcal{F}'$, then G_{i+1} is removed from $\mathcal{G}^i$ and nothing happens to $\mathcal{H}_1, \mathcal{H}_2$ so the claim is true for $i + 1$. If $G_{i+1} \in \mathcal{F} \cap \mathcal{F}'$, then, due to the fixed ordering of $2^{2^{[n]}}$, the antichain $\mathcal{A}^{i+1}$ is defined the same for both runs. This immediately yields that $\mathcal{G}^{i+1}$ is defined the same for both runs. Also, as for any $A \in \mathcal{A}^{i+1}$ either A becomes a set of $\mathcal{H}_1, \mathcal{H}_2$ in this step or never, and in the end these sets are the same, therefore they must be the same after step $i + 1$. $\square$

3 Proofs

3.1 $\vee_{r+1}$-free families and consequences - Theorem 7, 10, and 14

In this subsection we present the proofs of our theorems concerning $\vee_{r+1}$-free and $K_{s,1,t}$-free families. We restate the theorems here for convenience.

Theorem 7. *For any $s, t \in \mathbb{N}$ and $\varepsilon > 0$ there exist $n_0 = n_{\varepsilon,s,t}$ and $\delta > 0$ such that any $\mathcal{F} \subseteq 2^{[n]}$ of size at least $(2 + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ with $n \geq n_0$ contains at least $\delta n^{s+t}\binom{n}{\lfloor n/2 \rfloor}$ induced copies of $K_{s,1,t}$.*

Proof. Let $\mathcal{F} \subseteq 2^{[n]}$ be a family of sets of size $(2+\varepsilon)\binom{n}{\lfloor n/2 \rfloor}$. Let $\mathcal{D}$ be the family of those elements of $\mathcal{F}$ that are not the maximal element of an induced $\wedge_{\varepsilon n/10}$ and $\mathcal{U}$ the family of those that are not the minimal element of an induced $\vee_{\varepsilon n/10}$. By Theorem 18, we have $|\mathcal{D}|, |\mathcal{U}| \leq (1 + 4\varepsilon/10)\binom{n}{\lfloor n/2 \rfloor}$, thus $|\mathcal{F} \setminus (\mathcal{D} \cup \mathcal{U})| \geq \frac{\varepsilon}{5}\binom{n}{\lfloor n/2 \rfloor}$. Taking sets from $\mathcal{F} \setminus (\mathcal{D} \cup \mathcal{U})$ to play the role of the middle element of $K_{s,1,t}$, by definition of $\mathcal{D}$ and $\mathcal{U}$, we obtain at least $\frac{\varepsilon}{5}\binom{n}{\lfloor n/2 \rfloor}\binom{\varepsilon n/10}{s}\binom{\varepsilon n/10}{t}$ copies of $K_{s,1,t}$. $\square$

Theorem 10.
    (i) The number of induced $\vee_r$-free families is $2^{(1+o(1))\binom{n}{\lfloor n/2 \rfloor}}$.
    (ii) The number of induced $K_{s,1,t}$-free families in $2^{[n]}$ is $2^{(2+o(1))\binom{n}{\lfloor n/2 \rfloor}}$.

Proof. To prove (i), we apply our container lemma, Theorem 20 with $t = 1$. It shows that for every induced $\vee_{r+1}$-free family $\mathcal{F}$, there exist $\mathcal{H}_1, \mathcal{H}_2$ and $g(\mathcal{H}_1 \cup \mathcal{H}_2)$ such that $\mathcal{H}_1$ and $\mathcal{H}_2$ are disjoint, $|\mathcal{H}_1 \cup \mathcal{H}_2| \leq (r+1)\frac{3}{\varepsilon^2 n}\binom{n}{\lfloor n/2 \rfloor}$, $|g(\mathcal{H}_1 \cup \mathcal{H}_2)| \leq (1+\varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ and $\mathcal{F} \subseteq \mathcal{H}_1 \cup \mathcal{H}_2 \cup g(\mathcal{H}_1 \cup \mathcal{H}_2)$. Therefore, $|\mathcal{H}_1 \cup \mathcal{H}_2 \cup g(\mathcal{H}_1 \cup \mathcal{H}_2)| \leq (1+2\varepsilon)\binom{n}{\lfloor n/2 \rfloor}$. The number of subfamilies of such containers

is at most $2^{(1+2\varepsilon)\binom{n}{\lfloor n/2 \rfloor}}$, and the number of such containers is at most

$$\binom{2^n}{(r+1)\frac{3}{\varepsilon^2 n}\binom{n}{\lfloor n/2 \rfloor}} 2^{(r+1)\frac{3}{\varepsilon^2 n}\binom{n}{\lfloor n/2 \rfloor}}.$$

Indeed, the first term is an obvious upper bound on the number of families $\mathcal{H}_1 \cup \mathcal{H}_2$, and the second term is an obvious upper bound on the number of ways to partition $\mathcal{H}_1 \cup \mathcal{H}_2$ to $\mathcal{H}_1$ and $\mathcal{H}_2$. Using $\binom{n}{\lfloor n/2 \rfloor} = \Theta(\frac{1}{\sqrt{n}} 2^n)$ and $\binom{a}{b} \leq (\frac{ea}{b})^b$, we obtain that the number of induced $\vee_{r+1}$-free families in $2^{[n]}$ is at most

$$\binom{2^n}{(r+1)\frac{3}{\varepsilon^2 n}\binom{n}{\lfloor n/2 \rfloor}} 2^{(r+1)\frac{3}{\varepsilon^2 n}\binom{n}{\lfloor n/2 \rfloor}} 2^{(1+2\varepsilon)\binom{n}{\lfloor n/2 \rfloor}} = 2^{(1+2\varepsilon+O_{\varepsilon,r}(\frac{\log n}{n}))\binom{n}{\lfloor n/2 \rfloor}}.$$

The lower bound follows from the fact that every subfamily of the middle level is $\vee_r$-free.

To prove (ii), observe that every induced $K_{s,1,t}$-free family $\mathcal{F} \subseteq 2^{[n]}$ can be written as $\mathcal{F} = \mathcal{D} \cup \mathcal{U}$ such that $\mathcal{D}$ is induced $\wedge_s$-free and $\mathcal{U}$ is induced $\vee_t$-free. Indeed, let $\mathcal{D}$ be the family of those elements of $\mathcal{F}$ that are not the maximal element of an induced $\wedge_s$ and $\mathcal{U}$ be the family of those that are not the minimal element of an induced $\vee_t$. If $F \in \mathcal{F}$ does not belong to $\mathcal{D} \cup \mathcal{U}$, then F with its s subsets and its t supersets form an induced $K_{s,1,t}$, which contradicts the $K_{s,1,t}$-free property of $\mathcal{F}$. By part (i), there are $2^{(1+o(1))\binom{n}{\lfloor n/2 \rfloor}}$ induced $\vee_t$-free families in $2^{[n]}$ and there are $2^{(1+o(1))\binom{n}{\lfloor n/2 \rfloor}}$ induced $\wedge_s$-free families in $2^{[n]}$. Therefore there are at most $2^{(2+o(1))\binom{n}{\lfloor n/2 \rfloor}}$ induced $K_{s,1,t}$-free families in $2^{[n]}$. The lower bound immediately follows from the fact that every subfamily of the middle two levels is $K_{s,1,t}$-free. $\qquad\square$

Theorem 14. *If $p = \omega(1/n)$, then the following are true.*
(i) For any integer $r \geq 0$, the largest induced $\vee_{r+1}$-free family in $\mathcal{P}(n,p)$ has size $(1+o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p.
(ii) For any pair $s,t \geq 1$ of integers, the largest induced $K_{s,1,t}$-free family in $\mathcal{P}(n,p)$ has size $(2+o(1))p\binom{n}{\lfloor n/2 \rfloor}$ w.h.p.

Proof. To prove (i), we again apply Theorem 20 with $t = 1$. All calculations are very close to those in [12], which in turn are almost the same as those in [1], we include them for sake of completeness.

It is enough to prove the statement for $\varepsilon < \frac{1}{4}$ and set $\varepsilon_1 = \varepsilon/4$. We will show that w.h.p. for every $\vee_{r+1}$-free family $\mathcal{F}$ in $2^{[n]}$ of size at least $(1+\varepsilon)p\binom{n}{\lfloor n/2 \rfloor}$, not all sets of $\mathcal{F}$ remain in $\mathcal{P}(n,p)$. For every such $\mathcal{F}$, Theorem 20 with ε_1 in the role of ε, gives us $\mathcal{H}_1 = \mathcal{H}_1(\mathcal{F}), \mathcal{H}_2 = \mathcal{H}_2(\mathcal{F})$ such that

- $\mathcal{H}_1 \in \binom{2^{[n]}}{\leq (r+1)2^n n^{-1.9}}$; therefore the number of possible $\mathcal{H}_1$'s is at most

$$\sum_{a \leq (r+1)2^n n^{-1.9}} \binom{2^n}{a}.$$

11

Clearly, we have $\mathbb{P}(\mathcal{H}_1 \subseteq \mathcal{P}(n,p)) = p^{|\mathcal{H}_1|}$.

- $\mathcal{H}_2 \in \binom{2^{[n]}}{\leq 3(r+1)\binom{n}{\lfloor n/2 \rfloor}/(\varepsilon_1^2 n)}$ and $\mathcal{H}_2 \subseteq f(\mathcal{H}_1) \in \binom{2^{[n]}}{(2+\varepsilon_1)\binom{n}{\lfloor n/2 \rfloor}}$, so for fixed $\mathcal{H}_1$ the number of possible $\mathcal{H}_2$'s is at most

$$\left| \binom{f(\mathcal{H}_1)}{\leq 3(r+1)\binom{n}{\lfloor n/2 \rfloor}/(\varepsilon_1^2 n)} \right| \leq \sum_{b \leq 3(r+1)\binom{n}{\lfloor n/2 \rfloor}/(\varepsilon_1^2 n)} \binom{3\binom{n}{\lfloor n/2 \rfloor}}{b}.$$

Also, $\mathbb{P}(\mathcal{H}_2 \subseteq \mathcal{P}(n,p)) = p^{|\mathcal{H}_2|}$.

- For fixed $\mathcal{H}_1$ and $\mathcal{H}_2$ the corresponding $\mathcal{F}$'s are all subfamilies of $\mathcal{H}_1 \cup \mathcal{H}_2 \cup g(\mathcal{H}_1 \cup \mathcal{H}_2)$ and contain $\mathcal{H}_1 \cup \mathcal{H}_2$.

  1. Let $\mathcal{E}_{\mathcal{H}_1,\mathcal{H}_2}$ be the event that there exists *any* $\mathcal{F}$ with $\mathcal{H}_1(\mathcal{F}) = \mathcal{H}_1$, $\mathcal{H}_2(\mathcal{F}) = \mathcal{H}_2$, and $|\mathcal{F}| \geq (1+\varepsilon)p\binom{n}{\lfloor n/2 \rfloor}$, and

  2. let $\mathcal{E}_{g(\mathcal{H}_1 \cup \mathcal{H}_2)}$ be the event that $|\mathcal{P}(n,p) \cap g(\mathcal{H}_1 \cup \mathcal{H}_2)| \geq (1+\varepsilon)p\binom{n}{\lfloor n/2 \rfloor} - |\mathcal{H}_1 \cup \mathcal{H}_2|$ holds.

  We bound the probability of the event $\mathcal{E}_{\mathcal{H}_1,\mathcal{H}_2}$ by the probability of the event $\mathcal{E}_{g(\mathcal{H}_1 \cup \mathcal{H}_2)}$. Note that

$$(1+\varepsilon)p\binom{n}{\lfloor n/2 \rfloor} - |\mathcal{H}_1 \cup \mathcal{H}_2| \geq (1+\varepsilon/2)p\binom{n}{\lfloor n/2 \rfloor}$$

and

$$|g(\mathcal{H}_1 \cup \mathcal{H}_2)| \leq (1+\varepsilon_1)\binom{n}{\lfloor n/2 \rfloor} \leq (1+\varepsilon/4)\binom{n}{\lfloor n/2 \rfloor}.$$

Therefore, $|\mathcal{P}(n,p) \cap g(\mathcal{H}_1 \cup \mathcal{H}_2)|$ is binomially distributed with

$$\mathbb{E}(|\mathcal{P}(n,p) \cap g(\mathcal{H}_1 \cup \mathcal{H}_2)|) \leq (1+\varepsilon/4)p\binom{n}{\lfloor n/2 \rfloor},$$

so by Chernoff's inequality we have

$$\mathbb{P}(\mathcal{E}_{g(\mathcal{H}_1,\mathcal{H}_2)}) \leq \mathbb{P}\left[|\mathcal{P}(n,p) \cap g(\mathcal{H}_1 \cup \mathcal{H}_2)| \geq (1+\varepsilon/2)p\binom{n}{\lfloor n/2 \rfloor} \right] \leq e^{-\varepsilon^2 p \binom{n}{\lfloor n/2 \rfloor}/100}.$$

Note that $\mathcal{H}_1$, $\mathcal{H}_2$, and $g(\mathcal{H}_1 \cup \mathcal{H}_2)$ are disjoint, so the three events that $\mathcal{H}_1 \subseteq \mathcal{P}(n,p)$, $\mathcal{H}_2 \subseteq \mathcal{P}(n,p)$ and $\mathcal{E}_{g(\mathcal{H}_1 \cup \mathcal{H}_2)}$ are independent. Hence the probability that for fixed $\mathcal{H}_1$ and $\mathcal{H}_2$ there is a corresponding large induced $\vee_{r+1}$-free family $\mathcal{F}$, is at most $p^{|\mathcal{H}_1|+|\mathcal{H}_2|}e^{-\varepsilon^2 p\binom{n}{\lfloor n/2 \rfloor}/100}$. Summing up for all possible $\mathcal{H}_1$ and $\mathcal{H}_2$, we obtain that the probability Π, that there is an induced $\vee_{r+1}$-free family in $\mathcal{P}(n,p)$ of size $(1+\varepsilon)p\binom{n}{\lfloor n/2 \rfloor}$, is at most

$$\sum_{0 \leq a \leq (r+1)n^{-1.9}2^n} \sum_{0 \leq b \leq 3(r+1)\binom{n}{\lfloor n/2 \rfloor}/(\varepsilon_1^2 n)} \binom{2^n}{a} p^a \binom{3\binom{n}{\lfloor n/2 \rfloor}}{b} p^b e^{-\varepsilon^2 p\binom{n}{\lfloor n/2 \rfloor}/100}.$$

12

It is not hard to verify that the largest summand in the above sum belongs to the largest possible values of a and b. Therefore the above expression is bounded from above by

$$(r+1)n^{-1.9}2^n \frac{3(r+1)\binom{n}{\lfloor n/2 \rfloor}}{\varepsilon_1^2 n}\left(\frac{2^n}{(r+1)n^{-1.9}2^n}\right)\left(\frac{3\binom{n}{\lfloor n/2 \rfloor}}{\frac{3(r+1)\binom{n}{\lfloor n/2 \rfloor}}{(\varepsilon_1^2 n)}}\right)\cdot e^{-\varepsilon^2 p\left(\binom{n}{\lfloor n/2 \rfloor}\right)/100}p^{(r+1)n^{-1.9}2^n}p^{3(r+1)\binom{n}{\lfloor n/2 \rfloor}/(\varepsilon^2 n)}.$$

Observe that

$$(r+1)n^{-1.9}2^n 3(r+1)\binom{n}{\lfloor n/2 \rfloor}/(\varepsilon_1^2 n) \leq e^{O(n)} \leq e^{\varepsilon^2 p\left(\binom{n}{\lfloor n/2 \rfloor}\right)/400}.$$

Also, using $\binom{n}{k} \leq (\frac{en}{k})^k$ and $p\binom{n}{\lfloor n/2 \rfloor} = \omega(n^{-1.5}2^n)$, we have

$$\left(\frac{2^n}{(r+1)n^{-1.9}2^n}\right)p^{(r+1)n^{-1.9}2^n} \leq (en^{1.9}p)^{(r+1)n^{-1.9}2^n} \leq e^{O(n^{-1.9}2^n \ln n)} \leq e^{\varepsilon^2 p\left(\binom{n}{\lfloor n/2 \rfloor}\right)/400}.$$

Finally, by the same reasoning we have

$$\left(\frac{3\binom{n}{\lfloor n/2 \rfloor}}{3(r+1)\binom{n}{\lfloor n/2 \rfloor}/(\varepsilon_1^2 n)}\right)p^{3(r+1)\binom{n}{\lfloor n/2 \rfloor}/(\varepsilon_1^2 n)} \leq (e\varepsilon_1^2 np)^{3(r+1)\binom{n}{\lfloor n/2 \rfloor}/(\varepsilon_1^2 n)}$$

$$\leq e^{3(r+1)\binom{n}{\lfloor n/2 \rfloor}p\frac{\ln(np)}{(\varepsilon^2 np)}} \leq e^{\varepsilon^2 p\left(\binom{n}{\lfloor n/2 \rfloor}\right)/400}.$$

Therefore, the probability Π is at most $e^{-\varepsilon^2 p\left(\binom{n}{\lfloor n/2 \rfloor}\right)/400} = o(1)$, as required.

To prove (ii), let $\mathcal{F} \subseteq \mathcal{P}(n,p)$ be an induced $K_{s,1,t}$-free family. As in the proof of Theorem 10 (ii), let us consider the partition $\mathcal{F} = \mathcal{D} \cup \mathcal{U}$ with $\mathcal{D} = \{F \in \mathcal{F} : \nexists F_1, F_2, \ldots, F_s \text{ an antichain with } F_i \subsetneq F\}$ and $\mathcal{U} = \{F \in \mathcal{F} : \nexists F_1, F_2, \ldots, F_t \text{ an antichain with } F_i \supsetneq F\}$. By part (i) of this theorem, both $\mathcal{D}$ and $\mathcal{U}$ are of size at most $(1+o(1))p\binom{n}{\lfloor n/2 \rfloor}$ and thus $|\mathcal{F}| \leq (2+o(1))p\binom{n}{\lfloor n/2 \rfloor}$. $\square$

3.2 Results on trees - Theorem 5 and 6

Lemma 21. *Let $\overrightarrow{T}$ be a directed tree with t edges that does not contain directed paths of length 2. Then there exists a $\delta > 0$ such that the following holds: if $f(m)$ tends to infinity with m and $\overrightarrow{G}_m$ is a directed graph on m vertices with $f(m) \cdot m$ edges, then $\overrightarrow{G}$ contains $\delta f(m)^t m$ copies of $\overrightarrow{T}$.*

Proof. It is a well-known fact that every undirected graph contains a cut (i.e., a partition of its vertices into two and the edges between the parts) that contains at least half of its edges. This fact easily implies that the vertex set of $\overrightarrow{G}_m$ can be partitioned into $A \cup B$ such that there exist $\frac{f(m)m}{4}$ edges pointing from A to B. Let $\overrightarrow{H}_0$ denote this directed subgraph of $\overrightarrow{G}_m$. For

13

$i = 1, 2, \ldots, t-1$ we remove the set $U_i \subseteq V(\overrightarrow{H}_{i-1})$ of vertices that are adjacent to at most $\frac{f(m)}{8t}$ edges in $\overrightarrow{H}_{i-1}$. We also remove the set of adjacent edges to obtain $\overrightarrow{H}_i$. As for each i we remove at most $\frac{f(m)m}{8t}$ edges, $\overrightarrow{H}_{t-1}$ contains at least $\frac{f(m)m}{8}$ edges. Also, by definition, if $\overrightarrow{uv}$ is an edge in $\overrightarrow{H}_i$, then both u and v are adjacent to at least $\frac{f(m)}{8t}$ edges in $\overrightarrow{H}_{i-1}$. Based on these properties, we can embed $\overrightarrow{T}$ to $\overrightarrow{G}_m$ greedily as follows: we fix an ordering $e_1, e_2, \ldots, e_t$ of the edges of $\overrightarrow{T}$ such that $e_1, \ldots, e_j$ is a tree for every $j = 1, 2, \ldots, t$ and then embed e_i to an edge of $\overrightarrow{H}_{t-i}$. There are at least $\frac{f(m)m}{8}$ choices for the image of e_1 and at least $\frac{f(m)}{8t} - i$ choices for the image of e_i if $i \geq 2$. $\qquad\square$

Theorem 5. *Let T be any height 2 tree poset of $t+1$ elements. Then for any $\varepsilon > 0$ there exist $\delta > 0$ and n_0 such that for any $n \geq n_0$ any family $\mathcal{F} \subseteq 2^{[n]}$ of size $|\mathcal{F}| \geq (1+\varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta n^t \binom{n}{\lfloor n/2 \rfloor}$ copies of T.*

Proof. For a family $\mathcal{F} \subseteq 2^{[n]}$ of size at least $(1+\varepsilon)\binom{n}{\lfloor n/2 \rfloor}$, let us consider its directed comparability graph $\overrightarrow{G}_{\mathcal{F}}$. This is the graph with vertex set $\mathcal{F}$ where $\overrightarrow{FG}$ is an edge if and only if $F \subsetneq G$. By Theorem 2, $\overrightarrow{G}_{\mathcal{F}}$ contains at least $\varepsilon \frac{n}{4} \binom{n}{\lfloor n/2 \rfloor}$ edges and thus Lemma 21 can be applied with $m = \binom{n}{\lfloor n/2 \rfloor}$ and $f(m) = \varepsilon \frac{n}{4}$. $\qquad\square$

Proposition 22. *For any monotone tree poset T, we have $x(T) = |T| - 1 + \sum(h(T) - r(\ell))$, where $h(T)$ denotes the height of T, the summation is over all leaves ℓ of T, and $r(\ell)$ is the rank of ℓ, i.e., its distance from the root in the Hasse diagram of T plus 1.*

Proof. Let us generate an embedding i of T into the $h(T)$ middle levels. Assume without loss of generality that T is upward monotone, thus the root must be mapped to the lowest of the middle levels. We will refer to this level as the first middle level, and analogously the level above it is the second middle level, and the kth middle level is above the first level by $k-1$.

Observe that there are $(1 + o(1))\binom{n}{\lfloor n/2 \rfloor}$ ways to pick the image of the root. Then we define i by going through the elements of T in a non-decreasing order with respect to the rank $r(p)$. That means that for every element p, when we decide where to embed it, we have already embedded its *predecessor*, the (unique) element p' that is smaller than p such that there is no element p'' with $p' < p'' < p$.

If the predecessor p' of p is mapped to the jth middle level and we want to map p to the kth middle level, then the number of possibilities is $\Theta(n^{k-j})$. Let $f_i : T \to [h(T)]$ be a function defined by $f(p) = k$ if $i(p)$ is on the kth middle level. For a given $f : T \to [h(T)]$, the number of embeddings i with $i = f_i$ is $\Theta(n^{\sum_{p \in T}(f(p) - f(p'))})$, where p' is the predecessor of p and the summation goes over all elements of P apart from the root. Therefore, to obtain $x(T)$, we need to maximize $\Theta(n^{\sum_{p \in P}(f(p) - f(p'))})$. Clearly, when embedding a leaf ℓ, we must have $f(\ell) = h(T)$. Finally, observe that if for some non-leaf p, we have $f(p) > f(p') + 1$, then changing $f(p)$ to $f(p') + 1$ cannot decrease the sum (and strictly increases it, if p has at least two children). This

shows that a function f that maximizes the sum must satisfy $f(\ell) = h(T)$ for all leaves, and $f(p) = r(p)$ for all non-leaves. $\square$

Lemma 23. *Assume that the families $2^{[n]} \supset \mathcal{F}_1 \supset \mathcal{F}_2 \supset \cdots \supset \mathcal{F}_h$ satisfy the following properties for some $\delta_1, \delta_2, \delta_3 > 0$.*

i) $|\mathcal{F}_h| \geq \delta_1 \binom{n}{\lfloor n/2 \rfloor}$.

ii) For every $i = 2, 3, \ldots, h$ and $F \in \mathcal{F}_i$ there exist at least $\delta_2 n$ sets $F' \in \mathcal{F}_{i-1}$ with $F \subsetneq F'$.

iii) For every $i = 2, 3, \ldots, h$ and $F \in \mathcal{F}_i$ there exist at least $\delta_3 n^{i-1}$ sets $F' \in \mathcal{F}_1$ with $F \subsetneq F'$.

Then for any upward monotone tree poset T of height $h(T) = h$, $\mathcal{F}_1$ contains at least $\delta n^{x(T)} \binom{n}{\lfloor n/2 \rfloor}$ copies of T, where $\delta > 0$ is a constant depending on $\delta_1, \delta_2, \delta_3$ and $|T|$.

Proof. We generate embeddings of T as follows: we embed elements of T according to their rank. The root of T can be embedded to any set $F \in \mathcal{F}_h$. By i), we have at least $\delta_1 \binom{n}{\lfloor n/2 \rfloor}$ choices.

Any non-leaf element $x \in T$ with its predecessor embedded to $F_i \in \mathcal{F}_i$ can be embedded to any set $F' \in \mathcal{F}_{i-1}$ with $F_i \subsetneq F'$ that has not yet been used by the embedding. By ii), we have at least $\delta_2 n - |T|$ choices. If x is of rank $r(x)$ then this process will embed it to $\mathcal{F}_{h+1-r(x)}$.

If ℓ is a leaf vertex of rank $r(\ell)$ in T, then its predecessor is embedded into some $F \in \mathcal{F}_{h+2-r(\ell)}$. We can embed ℓ to any superset of F that has not yet been used. By iii), there exist at least $\delta_3 n^{h+1-r(\ell)} - |T|$ such sets.

Proposition 22 yields that the exponent of n in the number of embeddings generated this way will be exactly $x(T)$. (We get a factor of n for all vertices except for the root, and an additional $n^{h-r(\ell)}$ for leaves ℓ.) A copy corresponds to at most $|T|!$ embeddings. $\square$

Theorem 6. *For any monotone tree poset T and $\varepsilon > 0$, there exist $\delta > 0$ and n_0 such that for any $n \geq n_0$ any family $\mathcal{F} \subseteq 2^{[n]}$ of size $|\mathcal{F}| \geq (h(T) - 1 + \varepsilon) \binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta n^{x(T)} \binom{n}{\lfloor n/2 \rfloor}$ copies of T.*

Proof. Throughout the proof we can assume that n is sufficiently large. Let $h := h(T)$. To prove the theorem, we will find families $\mathcal{F} \supseteq \mathcal{F}_1 \supset \mathcal{F}_2 \supset \cdots \supset \mathcal{F}_h$ that satisfy the conditions of Lemma 23 for sufficiently small $(\delta_1, \delta_2, \delta_3)$ depending on ε and T. Let $\mathcal{F}_1 = \{G \in \mathcal{F} : ||G| - n/2| < n^{2/3}\}$. Then $|\mathcal{F} \setminus \mathcal{F}_1| \leq |\{G \subseteq [n] : ||G| - n/2| \geq n^{2/3}\}| \leq o\left(\binom{n}{\lfloor n/2 \rfloor}\right)$.

For $i = 2, 3, \ldots, h$ we define $\mathcal{F}_i$ as follows. Let $\varepsilon' := \varepsilon/2h$. A set $F \in \mathcal{F}_{i-1}$ is in $\mathcal{F}_i$ if the expected number of sets of $\mathcal{F}_{i-1}$ in a random chain between F and $[n]$ is at least $1 + \varepsilon'$. (F counts for all chains.)

Let $\mathcal{G}_i := \mathcal{F}_{i-1} \setminus \mathcal{F}_i$. We partition the $n!$ full chains from $\emptyset$ to $[n]$ as follows. Let $\mathbf{C}_G$ denote the collection of those chains whose smallest member from $\mathcal{G}_i$ is G. Let $\mathbf{C}_0$ denote the collection of those full chains that avoid $\mathcal{G}_i$. If we take a random full chain from any of these collections, the expected value of members of $\mathcal{G}_i$ in it is at most $1 + \varepsilon'$. Therefore the expected value taken

15

over all $n!$ full chains is also at most $1 + \varepsilon'$. This means that $\lambda_n(\mathcal{G}_i) := \sum_{F \in \mathcal{G}_i} \frac{1}{\binom{n}{|F|}} \leq 1 + \varepsilon'$ and therefore $|\mathcal{G}_i| \leq (1 + \varepsilon')\binom{n}{\lfloor n/2 \rfloor}$. These inequalities for $i = 2, 3, \ldots, h$ together imply i) of Lemma 23, if $\delta_1 < \varepsilon/2$.

Consider a set $F \in \mathcal{F}_{i-1}$. If F has at most $\delta_2 n$ supersets of size between $|F|+1$ and $|F|+h-1$ in $\mathcal{F}_{i-1}$ and at most $(\delta_2 + \delta_3)n^h$ supersets in all of $\mathcal{F}_{i-1}$, then the expected value of elements of $\mathcal{F}_{i-1}$ in a chain from F to $[n]$ is

$$\sum_{F \subsetneq F' \in \mathcal{F}_{i-1}} \frac{1}{\binom{n-|F|}{|F' \setminus F|}} \leq \frac{\delta_2 n}{n - |F|} + \frac{(\delta_2 + \delta_3)n^h}{\binom{n-|F|}{h}} \leq 3\delta_2 + 3^h h!(\delta_2 + \delta_3)$$

which is smaller than ε' if δ_2 and δ_3 are a sufficiently small positive numbers compared to h and ε'. This means that any $G \in \mathcal{F}_i$ has either at least $\delta_2 n$ proper supersets of size at most $|G|+h-1$ in $\mathcal{F}_{i-1}$ or at least $(\delta_2 + \delta_3)n^h$ supersets in all of $\mathcal{F}_{i-1}$.

The above statement trivially implies ii) of Lemma 23. We show that iii) of Lemma 23 follows as well. Let $F \in \mathcal{F}_i$ for some $2 \leq i \leq h$. We need to find at least $\delta_3 n^{i-1}$ supersets of F in $\mathcal{F}_1$. If it has $(\delta_2 + \delta_3)n^h$ supersets in $\mathcal{F}_{i-1}$ then we are obviously done, so assume that this is not the case. We define a directed graph as follows. F has at least $\delta_2 n$ supersets of size at most $|G| + h - 1$ in $\mathcal{F}_{i-1}$, let us direct an edge from F to all of them. Each of these sets will have at least $\delta_2 n$ supersets of size at most $|G| + 2(h-1)$ in $\mathcal{F}_{i-2}$, direct an edge from the subset to the supersets. Continue this operation until we reach $\mathcal{F}_1$. Now we have at least $(\delta_2 n)^{i-1}$ directed paths from G to members of $\mathcal{F}_1$. All these members are of size at most $|G| + (i-1)(h-1)$, therefore the number of paths leading to a single one of them is at most $((i-1)(h-1))^{i-1} < h^{2h}$. Thus we found at least $\frac{(\delta_2 n)^{i-1}}{h^{2h}}$ supersets of G in $\mathcal{F}_1$. This is more than $\delta_3 n^{i-1}$ if δ_3 is sufficiently small compared to δ_2 and h. $\qquad\square$

3.3 Diamond-free families - Theorem 8

Let us recall that $m_s = \lceil \log_2(s+2) \rceil$ and $m_s^* = \min\{m : s \leq \binom{m}{\lceil m/2 \rceil}\}$ and that for any integer $s \geq 2$, we have $x(D_s) = m_s$ and $x^*(D_s) = m_s^*$.

Theorem 8.

(i) If $s \in [2^{m_s-1} - 1, 2^{m_s} - \binom{m_s}{\lceil \frac{m_s}{2} \rceil} - 1]$, then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that every $\mathcal{F} \subseteq 2^{[n]}$ with $|\mathcal{F}| \geq (m_s + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta \cdot n^{m_s - 0.5}\binom{n}{\lfloor n/2 \rfloor}$ copies of D_s.

(ii) For any $\varepsilon > 0$ there exists a $\delta > 0$ such that every $\mathcal{F} \subseteq 2^{[n]}$ with $|\mathcal{F}| \geq (4 + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta \cdot n^{3.5}\binom{n}{\lfloor n/2 \rfloor}$ induced copies of D_4.

(iii) For any constant c with $1/2 < c < 1$ there exists an integer s_c such that if $s \geq s_c$ and $s \leq c\binom{m_s^*}{\lfloor m_s^*/2 \rfloor}$, then the following holds: for any $\varepsilon > 0$ there exists a $\delta > 0$ such that every $\mathcal{F} \subseteq 2^{[n]}$ with $|\mathcal{F}| \geq (m_s^* + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$ contains at least $\delta \cdot n^{m_s^* - 0.5}\binom{n}{\lfloor n/2 \rfloor}$ induced copies of D_s.

Proof. To prove (i), let us fix $\varepsilon > 0$ and let $\mathcal{F} \subseteq 2^{[n]}$ such that $|\mathcal{F}| \geq (m_s + \varepsilon)\binom{n}{\lfloor n/2 \rfloor}$. Let ℓ be the maximum integer such that $\binom{n}{\ell} \leq \frac{\varepsilon}{8\sqrt{n}}\binom{n}{\lfloor n/2 \rfloor}$. Then by Lemma 15, we have $|\binom{[n]}{\leq \ell} \cup \binom{[n]}{\geq n-\ell}| \leq \frac{\varepsilon}{2}\binom{n}{\lfloor n/2 \rfloor}$ and it is easy to verify that $\binom{n}{\ell} \geq \frac{\varepsilon}{10\sqrt{n}}\binom{n}{\lfloor n/2 \rfloor}$ if n is large enough. This implies that

1. $\mathcal{F}' = \mathcal{F} \setminus (\binom{[n]}{\leq \ell} \cup \binom{[n]}{\geq n-\ell})$ is of size at least $(m_s + \varepsilon/2)\binom{n}{\lfloor n/2 \rfloor}$, in particular $\sum_{F \in \mathcal{F}} |F|!(n - |F|)! \geq (m_s + \varepsilon/2)n!$;

2. for any $F \in \mathcal{F}'$, we have $|F|!(n - |F|)! \leq \frac{10}{\varepsilon}\sqrt{n}\lfloor n/2 \rfloor!\lceil n/2 \rceil!$.

We are going to count pairs $(F, \mathcal{C})$ with $F \in \mathcal{F}' \cap \mathcal{C}$ and $\mathcal{C}$ is a maximal chain in $[n]$. On the one hand, using the first point above, this is clearly $\sum_{F \in \mathcal{F}} |F|!(n - |F|)! \geq (m_s + \varepsilon/2)n!$. To count the pairs in another way, we will use the min-max partition of the maximal chains. Let $\mathbf{C}_n$ denote the set of all $n!$ maximal chains in $[n]$, and let us partition $\mathbf{C}_n$ into $\cup_{F,F'}\mathbf{C}_{F,F'}$, where F, F' run through all pairs $F \subseteq F'$ in $\mathcal{F}'$ and

$$\mathbf{C}_{F,F'} := \{\mathcal{C} \in \mathbf{C}_n : F \text{ is minimal in } \mathcal{F}' \cap \mathcal{C}, F' \text{ is maximal in } \mathcal{F}' \cap \mathcal{C}\}.$$

Those maximal chains that do not contain any $F \in \mathcal{F}'$ are gathered in $\mathbf{C}_\emptyset$. For any pair $F \subseteq F'$ in $\mathcal{F}'$ let us write $b(F, F') = |\mathcal{F}' \cap \{G : F \subsetneq G \subsetneq F'\}|$. Observe that the number of copies of D_s in $\mathcal{F}'$ is at least $\sum_{F,F'}\binom{b(F,F')}{s}$, in particular this number is at least $|\{(F, F') : b(F, F') \geq s\}|$. Finally, let $\mathbf{C}_j = \bigcup_{b(F,F')=j}\mathbf{C}_{F,F'}$ and $\mathbf{C}_{<s} = \mathbf{C}_\emptyset \cup \bigcup_{j=0}^{s-1}\mathbf{C}_j$ and $\mathbf{C}_{\geq s} = \mathbf{C}_n \setminus \mathbf{C}_{<s}$.

As $b(F, F') < s$ is equivalent to $[F, F'] \cap \mathcal{F}'$ being D_s-free, using Lemma 16, we obtain that the number of pairs $(F, \mathcal{C})$ with $F \in \mathcal{F}' \cap \mathcal{C}$ and $\mathcal{C} \in \mathbf{C}_{<s}$ is at most $m_s n!$. As a consequence, the number of pairs $(F, \mathcal{C})$ with $F \in \mathcal{F}' \cap \mathcal{C}$ and $\mathcal{C} \in \mathbf{C}_{\geq s}$ is at least $\frac{\varepsilon}{2}n!$. Observe that $|\mathbf{C}_{F,F'}| \leq |F|!(|F'| - |F|)!(n - |F'|)!$ and thus the number of pairs $(F'', \mathcal{C})$ with $F'' \in \mathcal{F}' \cap \mathcal{C}$ and $\mathcal{C} \in \mathbf{C}_{F,F'}$ is at most $(|F'| - |F| + 1)|F|!(|F'| - |F|)!(n - |F'|)!$. Also if $b(F, F') \geq s$, then $|F'| - |F| \geq m_s$. Fixing $|F|$ and using $\ell \leq |F|, |F'| \leq n - \ell$, it is easy to see that $(|F'| - |F| + 1)|F|!(|F'| - |F|)!(n - |F'|)!$ is convex in $|F'|$ and thus it is maximized either at $|F'| - |F| = m_s$ or when $|F'| = n - \ell$. Plugging in one obtains that the maximum is taken when $|F'| - |F| = m_s$. Using 2. from above, we obtain that for one fixed $\mathbf{C}_{F,F'} \subseteq \mathbf{C}_{\geq s}$ the number of pairs $(F'', \mathcal{C})$ with $F'' \in \mathcal{F}' \cap \mathcal{C}$ and $\mathcal{C} \in \mathbf{C}_{F,F'}$ is at most

$$(m_s+1)m_s!|F|!(n-|F|-m_s)! \leq (m_s+1)m_s!\frac{4^{m_s}}{n^{m_s}}|F|!(n-|F|)! \leq (m_s+1)m_s!\frac{4^{m_s}10}{\varepsilon n^{m_s-0.5}}\lfloor n/2 \rfloor!\lceil n/2 \rceil!.$$

Therefore the number of pairs F, F' with $b(F, F') \geq s$ and thus the number of copies of D_s in $\mathcal{F}'$ is at least

$$\frac{\varepsilon^2}{20(m_s + 1)!4^{m_s}}n^{m_s-0.5}\binom{n}{\lfloor n/2 \rfloor}$$

as claimed.

The proofs of (ii) and (iii) are basically the same. Instead of $b(F, F')$ one introduces $a(F, F')$, which is the maximum size of an antichain in $\mathcal{F}' \cap [F, F']$, and partitions $\mathbf{C}_n$ into $\mathbf{C}_{<s}$ and $\mathbf{C}_{\geq s}$

according to $a(F, F')$. If $a(F, F') < s$, then $\mathcal{F}' \cap [F, F']$ is induced D_s-free, so for (ii), instead of Lemma 16, one applies Lemma 17 (i), and for (iii), one applies Lemma 17 (ii) to $\mathbf{C}_{<s}$. The number of induced copies of D_s is at least the number of pairs with $a(F, F') \geq s$ and the computation to obtain a lower bound for this number is the same as in (i). $\qquad \square$

References

[1] J. Balogh, R. Mycroft, and A. Treglown. A random version of Sperner's theorem. *Journal of Combinatorial Theory, Series A*, 128:104–110, 2014.

[2] J. Balogh and A. Z. Wagner. Kleitman's conjecture about families of given size minimizing the number of k-chains. *Advances in Mathematics*, 330:229–252, 2018.

[3] B. Bukh. Set families with a forbidden subposet. *The Electronic Journal of Combinatorics*, 16(1):R142, 2009.

[4] T. Carroll and G. O. H. Katona. Bounds on maximal families of sets not containing three sets with $A \cap B \subseteq C$, $A \nsubseteq B$. *Order*, 25(3):229–236, 2008.

[5] S. Das, W. Gan, and B. Sudakov. Sperner's theorem and a problem of Erdős, Katona and Kleitman. *Combinatorics, Probability and Computing*, 24(4):585–608, 2015.

[6] A. P. Dove, J. R. Griggs, R. Kang, and J.-S. Sereni. Supersaturation in the Boolean lattice. *Integers: Electronic Journal of Combinatorial Number Theory*, 14:1–7, 2014.

[7] P. Erdős. On a lemma of Littlewood and Offord. *Bulletin of the American Mathematical Society*, 51(12):898–902, 1945.

[8] D. Gerbner and B. Patkós. *Extremal finite set theory*. CRC Press, 2018.

[9] J. R. Griggs and W.-T. Li. The partition method for poset-free families. *Journal of Combinatorial Optimization*, 25(4):587–596, 2013.

[10] J. R. Griggs, W.-T. Li, and L. Lu. Diamond-free families. *Journal of Combinatorial Theory, Series A*, 119(2):310–322, 2012.

[11] J. R. Griggs and L. Lu. On families of subsets with a forbidden subposet. *Combinatorics, Probability and Computing*, 18(5):731–748, 2009.

[12] K. A. Hogenson. *Random and deterministic versions of extremal poset problems*. PhD thesis, Iowa State University, 2016.

[13] G. O. H. Katona and T. G. Tarján. Extremal problems with excluded subgraphs in the n-cube. In *Graph Theory*, pages 84–93. Springer, 1983.

[14] D. Kleitman. A conjecture of Erdős-Katona on commensurable pairs among subsets of an n-set. In *Theory of Graphs, Proceedings of Colloquium held at Tihany, Hungary*, pages 187–207, 1966.

[15] D. Kleitman. On Dedekind's problem: the number of monotone Boolean functions. *Proceedings of the American Mathematical Society*, 21(3):677–682, 1969.

[16] A. D. Korshunov. The number of monotone Boolean functions (in Russian). *Problemy Kibernetiki*, 38:5–108, 1981.

[17] M. C. Neto and R. Morris. Maximum-size antichains in random set-systems. *Random Struct. Algorithms*, 49(2):308–321, 2016.

[18] B. Patkós. Induced and non-induced forbidden subposet problems. *Electronic Journal of Combinatorics*, 22(1):P1–30, 2015.

[19] B. Patkós. Supersaturation and stability for forbidden subposet problems. *Journal of Combinatorial Theory, Series A*, 136:220–237, 2015.

[20] W. Samotij. Subsets of posets minimising the number of chains. *Transactions of the American Mathematical Society*, 371(10):7259–7274, 2019.

[21] E. Sperner. Ein Satz über Untermengen einer endlichen Menge. *Mathematische Zeitschrift*, 27(1):544–548, 1928.