On families of subsets with a forbidden subposet

Jerrold R. Griggs* Linyuan Lu †

June 7, 2008

Abstract

Let $\mathcal{F} \subset 2^{[n]}$ be a family of subsets of $\{1, 2, \ldots, n\}$. For any poset H, we say $\mathcal{F}$ is H-free if $\mathcal{F}$ does not contain any subposet isomorphic to H. Katona and others have investigated the behavior of $\mathrm{La}(n, H)$, which denotes the maximum size of H-free families $\mathcal{F} \subset 2^{[n]}$. Here we use a new approach, which is to apply methods from extremal graph theory and probability theory to identify new classes of posets H, for which $\mathrm{La}(n, H)$ can be determined asymptotically as $n \to \infty$ for various posets H, including two-end-forks, up-down trees, and cycles C_{4k} on two levels.

Dedicated to Prof. William T. Trotter on the occasion of his 65th birthday

1 Introduction and Results

A poset $(S, \leq)$ is a set S equipped with a partial ordering $\leq$. We say a poset $(S, \leq)$ contains another poset $(S', \leq')$ if there exists an injection $f \colon S' \to S$, which preserves the partial ordering, meaning that whenever $u, v \in S'$ satisfy $u \leq' v$, we have $f(u) \leq f(v)$. In this case, S' is called a subposet of S.

Let $\mathcal{F} \subset 2^{[n]}$ be a family of subsets of $[n] := \{1, 2, \ldots, n\}$. For any poset H, we say $\mathcal{F}$ is H-free if the poset $(\mathcal{F}, \subseteq)$ does not contain H as a subposet. Let $\mathrm{La}(n, H)$ denote the largest size of H-free family of subsets of $[n]$. The fundamental result of this kind is for H being a chain P_2 of two elements. A P_2-free family is an antichain, and Sperner's Theorem [10] from 1928 gives us that $\mathrm{La}(n, P_2) = \binom{n}{\lfloor \frac{n}{2} \rfloor}$. For small posets H in general, it is interesting to compare $\mathrm{La}(n, H)$ to $\binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Erdős [5] extended Sperner's Theorem in 1945 to determine that $\mathrm{La}(n, P_k)$, where P_k is a chain (path) of k elements, is the sum of the $k-1$ middle binomial coefficients in n.

*Department of Mathematics, University of South Carolina, Columbia, SC 29208 USA (griggs@math.sc.edu). This author's work was supported in part by the conference "New Directions in Algorithms, Combinatorics, and Optimization," held in honor of Tom Trotter at Georgia Tech in May, 2008.

†Department of Mathematics, University of South Carolina, Columbia, SC 29208 USA (lu@math.sc.edu). This author was supported in part by NSF grant DMS 0701111.

Consequently, $\text{La}(n, P_k) \sim (k-1)\binom{n}{\lfloor \frac{n}{2} \rfloor}$, as $n \to \infty$. Let $h(P)$ denote the height of poset P, which is the largest cardinality of any chain in H. We are interested in the asymptotic behavior of $\text{La}(n, H)$ for other posets H of *height* k.

There have been several investigations already of height two posets. Thanh [11] extended Sperner's Theorem by showing that for all r, $\text{La}(n, V_r) \sim \binom{n}{\lfloor \frac{n}{2} \rfloor}$, where V_r is the r-fork, the height two poset with one element at the bottom level below each of r elements at the top level. (Especially, V_1 is P_2, while V_2 looks like the letter V.) It is important to note that we are not only excluding "induced" copies of a forbidden subposet H, e.g., V_3 is a subposet of P_4, so excluding V_3 subposets also excludes P_4.

DeBonis and Katona [3] determined that $\text{La}(n, B)$, where B is the Butterfly poset on four elements A_1, A_2, B_1, B_2 with each $A_1, A_2 \leq B_1, B_2$, is the sum of the two middle binomial coefficients in n. More generally, consider excluding the height two poset which is called (using graph-theoretic terminology) $K_{r,s}$, which has elements $A_i, 1 \leq i \leq r$ at the bottom level, elements $B_j, 1 \leq j \leq s$ at the top level, and for all i, j, $A_i \leq B_j$. DeBonis and Katona[3] extend the asymptotics for the butterfly B and show that $\text{La}(n, K_{r,s}) \sim 2\binom{n}{\lfloor \frac{n}{2} \rfloor}$ for all $r, s \geq 2$. Griggs and Katona [9] considered whether the asymptotics of excluding the N poset on four elements A_1, A_2, B_1, B_2 with $A_1 \leq B_1, A_2 \leq B_1, A_2 \leq B_2$ is similar to excluding V_2 or B. It turns out to be the former: $\text{La}(n, N) \sim \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

One new class of posets considered here we call a *baton* $P_k(s, t)$, which is a path P_k on k elements, $k \geq 3$, such that the bottom element is replicated $s-1$ times and the top element is replicated $t-1$ times, $s, t \geq 1$. That is, we have a height k poset with s (resp. t) independent elements on the bottom (resp., top) level. The particular case $P_k(1, r)$ (which resembles a palm tree), known as an r-fork with a k-shaft, has been examined by Katona and De Bonis [3]. They show

$$La(n, P_k(1,r)) \geq \sum_{i=\lfloor \frac{n-(k-2)}{2} \rfloor}^{\lfloor \frac{n+(k-2)}{2} \rfloor} \binom{n}{i} + \binom{n}{\lfloor \frac{n+k+1}{2} \rfloor} \left(\frac{r-1}{n} + \Omega(\frac{1}{n^2}) \right) \tag{1}$$

$$La(n, P_k(1,r)) \leq \sum_{i=\lfloor \frac{n-(k-2)}{2} \rfloor}^{\lfloor \frac{n+(k-2)}{2} \rfloor} \binom{n}{i} + \binom{n}{\lfloor \frac{n+k+1}{2} \rfloor} \left(\frac{z(k)+2(r-1)}{n} + \Omega(\frac{1}{n^2}) \right) \tag{2}$$

where $z(k) = \lfloor \frac{k^2}{2} \rfloor$ if $n+k$ is even and $z(k) = \lfloor \frac{(k-1)^2}{2} \rfloor$ if $n+k$ is odd.

The previously known maximum sizes of families of subsets of $[n]$ without a given pattern are listed in the following table.

Name	H	$\mathrm{La}(n,H)$	Reference
Chain P_r	$A_1 \subset \cdots \subset A_r$	$(r-1+o_n(1))\binom{[n]}{\lfloor \frac{n}{2}\rfloor}$	[5]
Butterfly B	$A_i \subset B_j$, for $1 \le i,j \le 2$	$(2+o_n(1))\binom{[n]}{\lfloor \frac{n}{2}\rfloor}$	[4]
$K_{r,s}$ $(r,s \ge 2)$	$A_i \subset B_j$, for $1 \le i \le r$, $1 \le j \le s$	$(2+o_n(1))\binom{[n]}{\lfloor \frac{n}{2}\rfloor}$	[3]
"N"	$A \subset B$, $C \subset B$, and $C \subset D$	$(1+o_n(1))\binom{[n]}{\lfloor \frac{n}{2}\rfloor}$	[9]
"V_r"	$A \subset B_i$, for $i=1,2,\ldots,r$	$(1+o_n(1))\binom{[n]}{\lfloor \frac{n}{2}\rfloor}$	[11]
${}_kV_r$	$A_1 \subset \cdots \subset A_k \subset B_i$, for $i=1,2,\ldots,r$	$(k+o_n(1))\binom{[n]}{\lfloor \frac{n}{2}\rfloor}$	[3]

Table 1: Previously known results in the literature

In this paper we give new asymptotic upper bounds on $\mathrm{La}(n,H)/\binom{n}{\lfloor \frac{n}{2}\rfloor}$ for several classes of posets H, and identify some new ones for which this ratio goes to 1 as $n \to \infty$. We first "roughly unify" the previous results on forks ${}_kV_r$ and on complete two level posets $K_{s,t}$ by considering batons $P_k(s,t)$. Note that the summation term in the bound, which appears repeatedly, is just the sum of the $k-1$ middle binomial coefficients in n.

Theorem 1 *For any $s,t \ge 1$ and $k \ge 3$, We have*

$$\mathrm{La}(n,P_k(s,t)) \le \sum_{i=\lfloor \frac{n-(k-2)}{2}\rfloor}^{\lfloor \frac{n+(k-2)}{2}\rfloor} \binom{n}{i} + \binom{n}{\lfloor \frac{n+k}{2}\rfloor}\left(\frac{2k(s+t-2)}{n} + O(n^{-3/2}\sqrt{\ln n})\right). \quad (3)$$

Consequently, as $n \to \infty$,

$$\mathrm{La}(n,P_k(s,t))/\binom{n}{\lfloor \frac{n}{2}\rfloor} \to k-1.$$

Remarks:

1. Theorem 1 (at $s=1$ and $t=r$) is better than inequality (2) for $k \ge 4r-3$. For small k and large r, inequality (2) gives a better constant in the second order term.

2. Note $\mathrm{La}(n,P_k(s,t)) \ge \mathrm{La}(n,P_k(1,\max\{s,t\}))$. ¿From inequality (1), we have

$$La(n,P_k(s,t)) \ge \sum_{i=\lfloor \frac{n-(k-2)}{2}\rfloor}^{\lfloor \frac{n+(k-2)}{2}\rfloor} \binom{n}{i} + \binom{n}{\lfloor \frac{n+k}{2}\rfloor}\left(\frac{\max\{s,t\}-1}{n} + \Omega(\frac{1}{n^2})\right). \quad (4)$$

   This lower bound (4) can be compared to the upper bound (3).

3. Note that $P_3(s,t)$ contains $P_2(s,t) = K_{s,t}$, the complete two level poset. Theorem 1 implies

$$\mathrm{La}(n,H) \le \left(2 + O(\frac{|H|}{n})\right)\binom{n}{\lfloor \frac{n}{2}\rfloor} \quad (5)$$

for all posets of height 2. The hidden constant in the second order term is slightly worse than that given in [3]. If H is not a subposet of the two middle layers of $2^{[n]}$ (for example H contains the butterfly B), then the equality in (5) holds.

An *up-down tree* T is a poset of height 2 that is also a tree as an undirected graph; its order is the number of elements, $|T|$.

Theorem 2 *For any up-down tree T with order t, we have*

$$\mathrm{La}(n,T) \leq \left(1 + \frac{16t}{n} + O\left(\frac{1}{n\sqrt{n \ln n}}\right)\right)\binom{n}{\lfloor \frac{n}{2} \rfloor}. \tag{6}$$

Consequently, as $n \to \infty$,

$$\mathrm{La}(n,T)/\binom{n}{\lfloor \frac{n}{2} \rfloor} \to 1.$$

After discovering the results above for batons and for up-down trees, we learned of new progress by Boris Bukh [1] that describes the asymptotic behavior of $\mathrm{La}(n,T)$ for every tree poset. Specifically, if T is any poset for which the Hasse diagram is a tree (connected and acyclic), then

$$\mathrm{La}(n,T) = (h(T)-1)\binom{n}{\lfloor \frac{n}{2} \rfloor}(1 + O(1/n)). \tag{7}$$

This implies the leading asymptotic behavior for batons and up-down trees in Theorems 1 and 2 above, though the proofs and error terms are different.

The butterfly poset B has been solved asymptotically, so it is next interesting to consider more generally the crowns $\mathcal{O}_{2k}$, which is the poset of height 2 that is a cycle of length $2k$ as an undirected graph. Of course, $\mathcal{O}_4$ is the butterfly poset, while $\mathcal{O}_6$ is noteworthy for being the middle two levels of the Boolean lattice B_3. We have the following theorem for crowns:

Theorem 3 *For $k \geq 2$, we have*

$$\begin{aligned} \mathrm{La}(n,\mathcal{O}_{4k}) &= (1+o_n(1))\binom{n}{\lfloor \frac{n}{2} \rfloor} && (8)\\ \mathrm{La}(n,\mathcal{O}_{4k-2}) &\leq \left(1 + \frac{\sqrt{2}}{2} + o_n(1)\right)\binom{n}{\lfloor \frac{n}{2} \rfloor}. && (9) \end{aligned}$$

So we see that the crowns $\mathcal{O}_{2k}$, $k \geq 3$, have $\mathrm{La}(n,\mathcal{O}_{2k})/\binom{n}{\lfloor \frac{n}{2} \rfloor}$ staying strictly below 2 asymptotically, unlike the Butterfly, the case $k=2$, where the ratio goes to 2. For even $k \geq 4$, the ratio goes to 1, while for odd $k \geq 3$ we only have an asymptotic upper bound.

The Theorem above for crowns is actually just a special case of the more general result which concerns a more general class of height 2 posets obtained from graphs in a natural way. The proof also relies on extremal graph theory. For a simple graph $G=(V,E)$, define a poset $P(G)$ on the set $V \cup E$ with the partial ordering $v < e$ if the edge e is incident at vertex v in G. For example, the crown poset $\mathcal{O}_{2k}$ is $P(G)$ when graph G is a k-cycle.

Theorem 4 *For any nonempty simple graph G with chromatic number $\chi(G)$, we have*

$$\mathrm{La}(n, P(G)) \le \left(1 + \sqrt{1 - \frac{1}{\chi(G) - 1}} + o_n(1)\right)\binom{n}{\lfloor \frac{n}{2} \rfloor}. \tag{10}$$

In particular, if G is a bipartite graph, then

$$\mathrm{La}(n, P(G)) = (1 + o_n(1))\binom{n}{\lfloor \frac{n}{2} \rfloor}. \tag{11}$$

Theorem 3 is a direct consequence of Theorem 4 by the observation $\mathcal{O}_{2k} = P(C_k)$.

In this theory we construct large families in the Boolean lattice that avoid a given subposet. This is analogous to the much-studied Turán theory of graphs, in which one seeks to maximize the number of edges on n vertices while avoiding a given subgraph. It is interesting that the theorem above applies the Turaán theory of graphs to give a useful bound in our ordered set theory.

The rest of the paper is organized as follows. Three probabilistic lemmas are given in Section 2, and the proofs of the theorems are given in section 3. We conclude with ideas for further research.

2 Lemmas

For any fixed poset H, $\mathrm{La}(n, H)$ is of magnitude $\Theta\left(\binom{n}{\lfloor \frac{n}{2} \rfloor}\right)$. The following lemma allows us to consider the families consisting only of subsets near the middle level.

Lemma 1 *For any positive integer n, we have*

$$\sum_{i > \frac{n}{2} + 2\sqrt{n \ln n}} \binom{n}{i} < \frac{2^n}{n^2}; \tag{12}$$

$$\sum_{i < \frac{n}{2} - 2\sqrt{n \ln n}} \binom{n}{i} < \frac{2^n}{n^2}. \tag{13}$$

Proof: Let $X_1, X_2, \ldots, X_n$ be n independent identically distributed $\{0, 1\}$ random variables with

$$\Pr(X_i = 0) = \Pr(X_i = 1) = \frac{1}{2}$$

for any $1 \le i \le n$. Apply Chernoff's inequality [2] to $X = \sum_{i=1}^n X_i$. We have

$$\Pr(X - E(X) > \lambda) < e^{-\frac{\lambda^2}{2n}}.$$

Choose $\lambda = 2\sqrt{n \ln n}$. We have

$$\begin{aligned} \sum_{i > \frac{n}{2} + 2\sqrt{n \ln n}} \binom{n}{i} 2^{-n} &= \Pr(X > \frac{n}{2} + \lambda) \\ &< e^{-\frac{\lambda^2}{2n}} \\ &= \frac{1}{n^2}. \end{aligned}$$

Inequality (12) has been proved. Inequality (13) is equivalent to inequality (12) by the symmetry of binomial coefficients $\binom{n}{i} = \binom{n}{n-i}$. $\square$

Apply Stirling's formula $n! = (1 + O(1/n)\sqrt{2\pi n})\frac{n^n}{e^n}$ to obtain the following approximation of $\binom{n}{\lfloor \frac{n}{2} \rfloor}$:

$$\begin{aligned} \binom{n}{\lfloor \frac{n}{2} \rfloor} &= \frac{n!}{\lfloor \frac{n}{2} \rfloor ! \lceil \frac{n}{2} \rceil !} \\ &= (1 + O(1/n)) \frac{\sqrt{2\pi n}\frac{n^n}{e^n}}{\sqrt{2\pi \lfloor \frac{n}{2} \rfloor} \frac{\lfloor \frac{n}{2} \rfloor^{\lfloor \frac{n}{2} \rfloor}}{e^{\lfloor \frac{n}{2} \rfloor}} \sqrt{2\pi \lceil \frac{n}{2} \rceil} \frac{\lceil \frac{n}{2} \rceil^{\lceil \frac{n}{2} \rceil}}{e^{\lceil \frac{n}{2} \rceil}}} \\ &= (1 + O(1/n)) \frac{\sqrt{2}}{\sqrt{\pi n}} 2^n. \end{aligned}$$

It implies that $\frac{2^n}{n^2} = (1 + O(1/n)) \frac{\sqrt{\pi/2}}{n^{3/2}} \binom{n}{\lfloor \frac{n}{2} \rfloor}$. For any family $\mathcal{F}$ of size $\Theta(\binom{n}{\lfloor \frac{n}{2} \rfloor})$, we can delete all subsets of sizes not in $(\frac{n}{2} - 2\sqrt{n \ln n}, \frac{n}{2} + 2\sqrt{n \ln n})$ from $\mathcal{F}$. We obtain a family of subsets that has about the same size of $\mathcal{F}$ and only contains subsets of sizes in $(\frac{n}{2} - 2\sqrt{n \ln n}, \frac{n}{2} + 2\sqrt{n \ln n})$.

Lemma 2 *Suppose X is a random variable which takes on nonnegative integer values. Let $f(x)$ and $g(x)$ be two nondecreasing functions defined for nonnegative integers x. Then*

$$\mathrm{E}(f(X)g(X)) \geq \mathrm{E}(f(X))\mathrm{E}(g(X)).$$

Proof: Apply the FKG inequality [8] over the totally ordered set of nonnegative integers. Alternately, here we give a simple direct proof.

For any integer $k \geq 1$, let h_k be the step function:

$$h_k(x) = \begin{cases} 0 & \text{if } 0 \leq x < k; \\ 1 & \text{if } x \geq k. \end{cases}$$

For integers $j \geq i \geq 1$, we observe that

$$\mathrm{E}(h_i(X)h_j(X)) \geq \mathrm{E}(h_i(X))\mathrm{E}(h_j(X)),$$

which holds sine

$$\begin{aligned}
\mathrm{E}(h_i(X)h_j(X)) &= \Pr(X \geq i \ \&\ X \geq j) \\
&= \Pr(X \geq j) \\
&\geq \Pr(X \geq i)\Pr(X \geq j) \\
&= \mathrm{E}(h_i(X))\mathrm{E}(h_j(X)).
\end{aligned}$$

We have

$$f(x) = f(0) + \sum_{k=1}^{\infty}(f(k) - f(k-1))h_k(x).$$

Similarly

$$g(x) = g(0) + \sum_{k=1}^{\infty}(g(k) - g(k-1))h_k(x).$$

All coefficients $f(k) - f(k-1)$ and $g(k) - g(k-1)$ are nonnegative. By linearity, we have

$$\begin{aligned}
\mathrm{E}(f(X)g(X)) &= \mathrm{E}((f(0) + \sum_{i=1}^{\infty}(f(i) - f(i-1))h_i(X))(g(0) + \sum_{j=1}^{\infty}(g(j) - g(j-1))h_j(X))) \\
&= f(0)g(0) + f(0)\sum_{j=1}^{\infty}(g(j) - g(j-1))\mathrm{E}(h_j(X)) \\
&\quad + g(0)\sum_{i=1}^{\infty}(f(i) - f(i-1))\mathrm{E}(h_i(X)) \\
&\quad + \sum_{i,j=1}^{\infty}(f(i) - f(i-1))(g(j) - g(j-1))\mathrm{E}(h_i(X)h_j(X)) \\
&\geq f(0)g(0) + f(0)\sum_{j=1}^{\infty}(g(j) - g(j-1))\mathrm{E}(h_j(X)) \\
&\quad + g(0)\sum_{i=1}^{\infty}(f(i) - f(i-1))\mathrm{E}(h_i(X)) \\
&\quad + \sum_{i,j=1}^{\infty}(f(i) - f(i-1))(g(j) - g(j-1))\mathrm{E}(h_i(X))\mathrm{E}(h_j(X)) \\
&= \mathrm{E}(f(X))\mathrm{E}(g(X)).
\end{aligned}$$

□

Lemma 3 *Suppose X is a random variable which takes on nonnegative integer values. For integers $k > r \geq 1$, if $\mathrm{E}(X) > k-1$, then*

$$\mathrm{E}\binom{X}{k} \geq \mathrm{E}\binom{X}{r}\frac{r!}{k!}\prod_{i=0}^{k-r-1}(\mathrm{E}(X) - r - i). \tag{14}$$

Proof: Define

$$f(x) = \begin{cases} \frac{r!}{k!}\prod_{i=0}^{k-r}(x-r-i) & \text{if } x > k-1 \\ 0 & \text{otherwise.} \end{cases}$$

and

$$g(x) = \begin{cases} \frac{1}{r!}\prod_{i=0}^{r-1}(x-i) & \text{if } x > r-1 \\ 0 & \text{otherwise.} \end{cases}.$$

Both $f(x)$ and $g(x)$ are nonnegative increasing functions. For each nonnegative integer x, we have $g(x) = \binom{x}{r}$ and $f(x)g(x) = \binom{x}{k}$. By applying Lemma 2 we obtain

$$\begin{aligned} \mathrm{E}\binom{X}{k} &= \mathrm{E}(f(X)g(X)) \\ &\geq \mathrm{E}(f(X))\mathrm{E}(g(X)) \\ &= \mathrm{E}(f(X))\mathrm{E}\binom{X}{r} \\ &\geq f(\mathrm{E}(X))\mathrm{E}\binom{X}{r}, \end{aligned}$$

where the last inequality follows from since $f(x)$ is concave upward. $\square$

3 Proofs of theorems

Proof of Theorem 1: We let $\epsilon = \frac{2k(s+t-2)}{\frac{n}{2}-2\sqrt{n\ln n}}$, and

$$f = f(n,k,s,t) = \sum_{i=\lfloor\frac{n-(k-2)}{2}\rfloor}^{\lfloor\frac{n+(k-2)}{2}\rfloor}\binom{n}{i} + \binom{n}{\lfloor\frac{n+k}{2}\rfloor}\epsilon.$$

Suppose $\mathcal{F}$ is a family of subsets of $[n]$ with $|\mathcal{F}| > f + \frac{2^{n+1}}{n^2}$. By removing all subsets of size outside $(\frac{n}{2} - 2\sqrt{n\ln n}, \frac{n}{2} + 2\sqrt{n\ln n})$, we can assume $\mathcal{F}$ only contains subsets of sizes in $(\frac{n}{2} - 2\sqrt{n\ln n}, \frac{n}{2} + 2\sqrt{n\ln n})$ and $|\mathcal{F}| > f$.

We would like to show $\mathcal{F}$ contains $P_k(s,t)$. We will prove this statement by contradiction. Suppose that $\mathcal{F}$ is $P_k(s,t)$-free. Take a random permutation $\sigma \in S_n$. Consider a random full (maximal) chain C_σ

$$\emptyset \subset \{\sigma_1\} \subset \{\sigma_1, \sigma_2\} \subset \cdots \subset \{\sigma_1, \sigma_2, \cdots, \sigma_n\}.$$

Let X be the random number counting $|\mathcal{F} \cap C_\sigma|$. On the one hand, we have

$$\begin{aligned} \mathrm{E}(X) &= \sum_{F\in\mathcal{F}} \frac{1}{\binom{n}{|F|}} && (15) \\ &> k-1+\epsilon, && (16) \end{aligned}$$

since the sum is minimized, for a family of subsets on $[n]$ of size f by taking the f sets closest to the middle size $n/2$, which means taking the $k-1$ middle levels and the remaining sets at the next closest level to the middle, $\lfloor \frac{n+k}{2} \rfloor$.

Apply Lemma 3 with $r = k-1$

$$\begin{aligned} \mathrm{E}\binom{X}{k} &\geq \frac{1}{k}\mathrm{E}\binom{X}{k-1}(\mathrm{E}(X) - k + 1) \\ &> \frac{\epsilon}{k}\mathrm{E}\binom{X}{k-1}. \end{aligned} \tag{17}$$

On the other hand, we will compute $\mathrm{E}\binom{X}{k}$ directly. By coumting chains, a subchain of length k in $\mathcal{F}$,

$$F_1 \subset F_2 \subset \cdots \subset F_k,$$

is in the random chain C_σ with probability

$$\frac{|F_1|!(|F_2| - |F_1|)! \cdots (n - |F_k|)!}{n!}.$$

By linearity, we have

$$\mathrm{E}\binom{X}{k} = \sum_{\substack{F_1,\ldots,F_k \in \mathcal{F} \\ F_1 \subset \cdots \subset F_k}} \frac{|F_1|!(|F_2| - |F_1|)! \cdots (n - |F_k|)!}{n!}. \tag{18}$$

We can rewrite equation (18) as

$$\mathrm{E}\binom{X}{k} = \sum_{\substack{F_2,\ldots,F_{k-1} \in \mathcal{F} \\ F_2 \subset \cdots \subset F_{k-1}}} \frac{|F_2|! \cdots (n - |F_{k-1}|)!}{n!} \sum_{\substack{F_1 \in \mathcal{F} \\ F_1 \subset F_2}} \frac{1}{\binom{|F_2|}{|F_1|}} \sum_{\substack{F_k \in \mathcal{F} \\ F_{k-1} \subset F_k}} \frac{1}{\binom{n-|F_{k-1}|}{n-|F_k|}}. \tag{19}$$

Since $\mathcal{F}$ is $P_k(s,t)$-free, for a fixed $F_2, \ldots, F_{k-1}$, either "the number of F_1 satisfying $F_1 \subset F_2$ is at most $s-1$" or "the number of F_k satisfying $F_{k-1} \subset F_k$ is at most $t-1$". Let $\mathcal{A}$ be the set of $k-2$-chains $F_2 \subset \ldots \subset F_{k-1}$ in $\mathcal{F}$ so that the number of $F_1 \in \mathcal{F}$, $F_1 \subset F_2$, is at most $s-1$. Let $\mathcal{B}$ be the set of $k-2$-chains $F_2 \subset \ldots \subset F_{k-1}$ in $\mathcal{F}$ so that the number of $F_k \in \mathcal{F}$, $F_{k-1} \subset F_k$, is at most $t-1$. The union of $\mathcal{A}$ and $\mathcal{B}$ covers all $k-2$-chains in $\mathcal{F}$. We have

$$\begin{aligned} \mathrm{E}\binom{X}{k} \leq &\sum_{(F_2,\ldots,F_{k-1}) \in \mathcal{A}} \frac{|F_2|! \cdots (n - |F_{k-1}|)!}{n!} \sum_{\substack{F_1 \in \mathcal{F} \\ F_1 \subset F_2}} \frac{1}{\binom{|F_2|}{|F_1|}} \sum_{\substack{F_k \in \mathcal{F} \\ F_{k-1} \subset F_k}} \frac{1}{\binom{n-|F_{k-1}|}{n-|F_k|}} \\ &+ \sum_{(F_2,\ldots,F_{k-1}) \in \mathcal{B}} \frac{|F_2|! \cdots (n - |F_{k-1}|)!}{n!} \sum_{\substack{F_1 \in \mathcal{F} \\ F_1 \subset F_2}} \frac{1}{\binom{|F_2|}{|F_1|}} \sum_{\substack{F_k \in \mathcal{F} \\ F_{k-1} \subset F_k}} \frac{1}{\binom{n-|F_{k-1}|}{n-|F_k|}}. \end{aligned} \tag{20}$$

For the summation over $\mathcal{A}$, the number of F_1 satisfying $F_1 \subset F_2$ is at most $s-1$. We have

$$\sum_{\substack{F_1 \in \mathcal{F} \\ F_1 \subset F_2}} \frac{1}{\binom{|F_2|}{|F_1|}} \leq \frac{(s-1)}{\frac{n}{2} - 2\sqrt{n \ln n}}. \tag{21}$$

Apply inequality (21) to the first summation in (20).

$$\begin{aligned}
&\sum_{(F_2,\ldots,F_{k-1})\in\mathcal{A}} \frac{|F_2|!\cdots(n-|F_{k-1}|)!}{n!} \sum_{\substack{F_1\in\mathcal{F}\\ F_1\subset F_2}} \frac{1}{\binom{|F_2|}{|F_1|}} \sum_{\substack{F_k\in\mathcal{F}\\ F_{k-1}\subset F_k}} \frac{1}{\binom{n-|F_{k-1}|}{n-|F_k|}} \\
\leq\ & \sum_{(F_2,\ldots,F_{k-1})\in\mathcal{A}} \frac{|F_2|!\cdots(n-|F_{k-1}|)!}{n!} \sum_{\substack{F_k\in\mathcal{F}\\ F_{k-1}\subset F_k}} \frac{1}{\binom{n-|F_{k-1}|}{n-|F_k|}} \frac{(s-1)}{\frac{n}{2}-2\sqrt{n\ln n}} \\
\leq\ & \sum_{\substack{F_2,\ldots,F_{k-1}\in\mathcal{F}\\ F_2\subset\cdots\subset F_{k-1}}} \frac{|F_2|!\cdots(n-|F_{k-1}|)!}{n!} \sum_{\substack{F_k\in\mathcal{F}\\ F_{k-1}\subset F_k}} \frac{1}{\binom{n-|F_{k-1}|}{n-|F_k|}} \frac{(s-1)}{\frac{n}{2}-2\sqrt{n\ln n}} \\
=\ & \mathrm{E}\binom{X}{k-1} \frac{(s-1)}{\frac{n}{2}-2\sqrt{n\ln n}}. \qquad (22)
\end{aligned}$$

For the summation over $\mathcal{B}$, the number of F_k satisfying $F_{k-1}\subset F_k$ is at most $t-1$. We have

$$\sum_{\substack{F_k\in\mathcal{F}\\ F_{k-1}\subset F_k}} \frac{1}{\binom{n-|F_{k-1}|}{n-|F_k|}} \leq \frac{(t-1)}{\frac{n}{2}-2\sqrt{n\ln n}}. \qquad (23)$$

An inequality similar to (22) can be obtained:

$$\sum_{(F_2,\ldots,F_{k-1})\in\mathcal{B}} \frac{|F_2|!\cdots(n-|F_{k-1}|)!}{n!} \sum_{\substack{F_1\in\mathcal{F}\\ F_1\subset F_2}} \frac{1}{\binom{|F_2|}{|F_1|}} \sum_{\substack{F_k\in\mathcal{F}\\ F_{k-1}\subset F_k}} \frac{1}{\binom{n-|F_{k-1}|}{n-|F_k|}} \leq \mathrm{E}\binom{X}{k-1} \frac{(t-1)}{\frac{n}{2}-2\sqrt{n\ln n}}. \qquad (24)$$

Combining inequalities (20), (22) and (24), we have

$$\mathrm{E}\binom{X}{k} \leq \mathrm{E}\binom{X}{k-1} \frac{s+t-2}{\frac{n}{2}-2\sqrt{n\ln n}}. \qquad (25)$$

¿From inequalities (17) and (25), and the fact that $\mathrm{E}\binom{X}{k-1} > 0$, we have

$$\frac{\epsilon}{k} < \frac{s+t-2}{\frac{n}{2}-2\sqrt{n\ln n}}$$

, which contradicts our choice of ϵ. □

Proof of Theorem 2: Let $\mathcal{F}$ be a T-free family of subsets of $[n]$. By removing at most $\frac{2^{n+1}}{n^2}$ subsets, without loss of generality, we can assume $\mathcal{F}$ consists of subsets of sizes in $(\frac{n}{2}-2\sqrt{n\ln n}, \frac{n}{2}+2\sqrt{n\ln n})$ and $|\mathcal{F}| > (1+\epsilon)\binom{n}{\lfloor \frac{n}{2}\rfloor}$. Here $\epsilon = \frac{2t}{n} + \frac{16t}{n\sqrt{n\ln n}}$.

Let X be the same variable as defined in the proof of Theorem 1. Recall

$$\mathrm{E}(X) = \sum_{F\in\mathcal{F}} \frac{1}{\binom{n}{|F|}}. \qquad (26)$$

We have

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}} \\
&\geq \frac{|\mathcal{F}|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \\
&> 1 + \epsilon. \qquad (27)
\end{aligned}$$

Using that the variance of X is nonnegative (or applying Lemma 3 with $r = 1$ and $k = 2$) we have

$$\mathrm{E}\binom{X}{2} \geq \frac{1}{2}\mathrm{E}(X)(\mathrm{E}(X) - 1). \qquad (28)$$

¿From inequality (27) and (28), we get

$$\mathrm{E}\binom{X}{2} > \frac{\epsilon}{2}\mathrm{E}(X). \qquad (29)$$

A simple case of inequality (18) with $k = 2$ is

$$\mathrm{E}\binom{X}{2} = \sum_{\substack{F_1, F_2 \in \mathcal{F} \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2| - |F_1|)!(n - |F_2|)!}{n!}. \qquad (30)$$

Now partition $\mathcal{F}$ into $\mathcal{A} \cup \mathcal{B}$ randomly. With probability $\frac{1}{4}$, a pair (F_1, F_2) has $F_1 \in \mathcal{A}$ and $F_2 \in \mathcal{B}$. There is a partition $\mathcal{F} = \mathcal{A} \cup \mathcal{B}$ satisfying

$$\sum_{\substack{F_1 \in \mathcal{A}, F_2 \in \mathcal{B} \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2| - |F_1|)!(n - |F_2|)!}{n!} > \frac{\epsilon}{8}\mathrm{E}(X). \qquad (31)$$

Now we consider an edge-weighted bipartite graph G with $V(G) = \mathcal{A} \cup \mathcal{B}$. such that F_1F_2 is an edge of G if $F_1 \in \mathcal{A}$, $F_2 \in \mathcal{B}$, and $F_1 \subset F_2$. Each edge F_1F_2 has weight $\frac{|F_1|!(|F_2|-|F_1|)!(n-|F_2|)!}{n!}$. Inequality (31) states that the total sum of edge-weights is greater than $\frac{\epsilon}{8}\mathrm{E}(X)$.

For any $F_1 \in \mathcal{A}$, the weighted degree of F_1 is

$$d_{F_1} = \frac{1}{\binom{n}{|F_1|}} \sum_{\substack{F_2 \in \mathcal{B} \\ F_1 \subset F_2}} \frac{1}{\binom{n-|F_1|}{n-|F_2|}}. \qquad (32)$$

Similarly, the weighted degree of $F_2 \in \mathcal{B}$ is

$$d_{F_2} = \frac{1}{\binom{n}{|F_2|}} \sum_{\substack{F_1 \in \mathcal{A} \\ F_1 \subset F_2}} \frac{1}{\binom{|F_2|}{|F_1|}}. \qquad (33)$$

We delete vertices F with weighted degree less than $\frac{\epsilon}{8}\frac{1}{\binom{n}{|F|}}$ recursively until all remaining vertices have weighted degree at least $\frac{\epsilon}{8}\frac{1}{\binom{n}{|F|}}$ in the remaining graph, call it G', which has vertex partition $\mathcal{A}' \cup \mathcal{B}'$ with $A' \subset A$ and $B' \subset B$. The sum of edge-weights in G' is at least

$$
\begin{aligned}
& \sum_{\substack{F_1 \in \mathcal{A}', F_2 \in \mathcal{B}' \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2|-|F_1|)!(n-|F_2|)!}{n!} \\
\geq & \sum_{\substack{F_1 \in \mathcal{A}, F_2 \in \mathcal{B} \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2|-|F_1|)!(n-|F_2|)!}{n!} \\
& - \sum_{\substack{F_1 \in \mathcal{A}\setminus\mathcal{A}', F_2 \in \mathcal{B} \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2|-|F_1|)!(n-|F_2|)!}{n!} \\
& - \sum_{\substack{F_1 \in \mathcal{A}, F_2 \in \mathcal{B}\setminus\mathcal{B}' \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2|-|F_1|)!(n-|F_2|)!}{n!} \\
= & \sum_{\substack{F_1 \in \mathcal{A}, F_2 \in \mathcal{B} \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2|-|F_1|)!(n-|F_2|)!}{n!} \\
& - \sum_{F_1 \in \mathcal{A}\setminus\mathcal{A}'} \frac{d_{F_1}}{\binom{n}{|F_1|}} - \sum_{F_2 \in \mathcal{B}\setminus\mathcal{B}'} \frac{d_{F_2}}{\binom{n}{|F_2|}} \\
> & \sum_{\substack{F_1 \in \mathcal{A}, F_2 \in \mathcal{B} \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2|-|F_1|)!(n-|F_2|)!}{n!} \\
& - \sum_{F_1 \in \mathcal{A}\setminus\mathcal{A}'} \frac{\epsilon}{8\binom{n}{|F_1|}} - \sum_{F_2 \in \mathcal{B}\setminus\mathcal{B}'} \frac{\epsilon}{8\binom{n}{|F_2|}} \\
\geq & \sum_{\substack{F_1 \in \mathcal{A}, F_2 \in \mathcal{B} \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2|-|F_1|)!(n-|F_2|)!}{n!} - \sum_{F \in \mathcal{F}} \frac{\epsilon}{8\binom{n}{|F|}} \\
\geq & \sum_{\substack{F_1 \in \mathcal{A}, F_2 \in \mathcal{B} \\ F_1 \subset F_2}} \frac{|F_1|!(|F_2|-|F_1|)!(n-|F_2|)!}{n!} - \frac{\epsilon}{8}\mathrm{E}(X).
\end{aligned}
$$

Since the last expression is positive by (31), both families $\mathcal{A}'$ and $\mathcal{B}'$ are non-empty.

By construction, every vertex in the remaining bipartite graph G' has weighted degree at least $\frac{\epsilon}{8}\frac{1}{\binom{n}{|F|}}$. For any $F_1 \in \mathcal{A}'$, by (32) we have

$$\sum_{\substack{F_2 \in \mathcal{B} \\ F_1 \subset F_2}} \frac{1}{\binom{n-|F_1|}{|F_2|-|F_1|}} \geq \frac{\epsilon}{8}. \tag{34}$$

Note

$$\binom{n-|F_1|}{|F_2|-|F_1|} \geq n-|F_1| \geq \frac{n}{2} - 2\sqrt{n\ln n}. \tag{35}$$

Combining inequalities (34) and (35), we have

$$\sum_{\substack{F_2\in\mathcal{B}'\\F_1\subset F_2}} 1 \geq \frac{\epsilon}{8}(\frac{n}{2} - 2\sqrt{n\ln n}). \tag{36}$$

Similarly, for any $F_1 \in \mathcal{A}'$,

$$\sum_{\substack{F_1\in\mathcal{A}'\\F_1\subset F_2}} 1 \geq \frac{\epsilon}{8}\left(\frac{n}{2} - 2\sqrt{n\ln n}\right). \tag{37}$$

In other words, the minimum degree (in the usual sense) of G' is at least $\frac{\epsilon}{8}(\frac{n}{2}-2\sqrt{n\ln n}) > t$ for the choice of ϵ.

A subgraph of G' which is isomorphic to T can be constructed as follows. For any $u \in V(T)$, map u to any vertex v of G'. Map the neighbors of u in T to the neighbors of v in G', and so on. Since the minimum degree is at least t, we can always find new vertex which has not been selected yet. This greedy algorithm finds a subposet isomorphic to T. □

Proof of Theorem 4: Let $\mathcal{F}$ be any $P(G)$-free subsets of $[n]$. By removing at most $\frac{2^{n+1}}{n^2}$ subsets, we can assume that $\mathcal{F}$ contains the subsets of sizes only in the interval $(\frac{n}{2}-2\sqrt{n\ln n}, \frac{n}{2}+2\sqrt{n\ln n})$. Let X be the random number defined in the proof of Theorem 1. We claim $\mathrm{E}(X) = 1 + o_n(1)$. Recall

$$\mathrm{E}(X) = \sum_{F\in\mathcal{F}} \frac{1}{\binom{n}{|F|}}, \tag{38}$$

so that $|\mathcal{F}| \leq E(X)\binom{n}{\lfloor\frac{n}{2}\rfloor}$. We obtain an upper bound on $E(X)$. As before, we have

$$\mathrm{E}\binom{X}{2} \geq \frac{1}{2}\mathrm{E}(X)(\mathrm{E}(X)-1). \tag{39}$$

We will bound $\mathrm{E}\binom{X}{2}$ in terms of $E(X)$. Recall

$$\mathrm{E}\binom{X}{2} = \sum_{\substack{A,B\in\mathcal{F}\\A\subset B}} \frac{|A|!(|B|-|A|)!(n-|B|)!}{n!}. \tag{40}$$

We split the summation into two parts, depending on whether $|B|-|A| = 1$ or $|B|-|A| > 1$.

For the case that $|B|-|A| > 1$, let Y be the random variable counting a triple (A,S,B) satisfying

$$A \subset S \subset B \quad A, B \in \mathcal{F}.$$

We have

$$\begin{aligned}
\mathrm{E}(Y) &= \sum_{\substack{A,B\in\mathcal{F},S\\ A\subset S\subset B}} \frac{|A|!(|S|-|A|)!(|B|-|S|)!(n-|B|)!}{n!} \\
&= \sum_{\substack{A,B\in\mathcal{F}\\ A\subset B}} \frac{|A|!(|B|-|A|)!(n-|B|)!}{n!} \sum_{S:A\subset S\subset B} \frac{1}{\binom{|B|-|A|}{|S|-|A|}} \\
&= \sum_{\substack{A,B\in\mathcal{F}\\ A\subset B}} \frac{|A|!(|B|-|A|)!(n-|B|)!}{n!}(|B|-|A|-1) \\
&\geq \sum_{\substack{A,B\in\mathcal{F}\\ A\subset B,|B|-|A|>1}} \frac{|A|!(|B|-|A|)!(n-|B|)!}{n!}. \qquad (41)
\end{aligned}$$

Denote the number of vertices in G by v and the number of edges in G by m. Since $\mathcal{F}$ is $P(G)$-free, there are no $v+m$ subsets $A_1, A_2, \ldots, A_v, B_1, \ldots, B_m \in \mathcal{F}$ satisfying $A_i \subset S \subset B_j$ for $1 \leq i \leq v$ and $1 \leq j \leq m$.

For any fixed subset S, either "at most $m-1$ subsets in $\mathcal{F}$ are supersets of S" or "at most $v-1$ subsets in $\mathcal{F}$ are subsets of S". Define

$$\mathcal{G}_1 = \{S \mid |S| \in (\frac{n}{2} - 2\sqrt{n\ln n}, \frac{n}{2} + 2\sqrt{n\ln n}), S \text{ has at most } v-1 \text{ subsets in } \mathcal{F}\}.$$

$$\mathcal{G}_2 = \{S \mid |S| \in (\frac{n}{2} - 2\sqrt{n\ln n}, \frac{n}{2} + 2\sqrt{n\ln n}), S \text{ has at most } m-1 \text{ supersets in } \mathcal{F}\}.$$

$\mathcal{G}_1 \cup \mathcal{G}_2$ covers all subsets with sizes in $(\frac{n}{2} - 2\sqrt{n\ln n}, \frac{n}{2} + 2\sqrt{n\ln n})$. Rewrite $\mathrm{E}(Y)$ as

$$\mathrm{E}(Y) = \sum_{S:||S|-\frac{n}{2}|<2\sqrt{n\ln n}} \frac{1}{\binom{n}{|S|}} \sum_{\substack{A\in\mathcal{F}\\ A\subset S}} \frac{1}{\binom{|S|}{|A|}} \sum_{\substack{B\in\mathcal{F}\\ S\subset B}} \frac{1}{\binom{n-|S|}{n-|B|}}. \qquad (42)$$

For $S \in \mathcal{G}_1$, we have

$$\sum_{B\in\mathcal{F},S\subset B} \frac{1}{\binom{n-|S|}{n-|B|}} \leq \frac{m-1}{\frac{n}{2} - 2\sqrt{n\ln n}}. \qquad (43)$$

It implies

$$\begin{aligned}
\sum_{S\in\mathcal{G}_2} \frac{1}{\binom{n}{|S|}} \sum_{\substack{A\in\mathcal{F}\\ A\subset S}} \frac{1}{\binom{|S|}{|A|}} \sum_{\substack{B\in\mathcal{F}\\ S\subset B}} \frac{1}{\binom{n-|S|}{n-|B|}} &\leq \sum_{S\in\mathcal{G}_1} \frac{1}{\binom{n}{|S|}} \sum_{\substack{A\in\mathcal{F}\\ A\subset S}} \frac{1}{\binom{|S|}{|A|}} \frac{m-1}{\frac{n}{2} - 2\sqrt{n\ln n}} \\
&\leq \mathrm{E}(X) 4\sqrt{n\ln n} \frac{m-1}{\frac{n}{2} - 2\sqrt{n\ln n}}. \qquad (44)
\end{aligned}$$

Similarly, we have

$$\sum_{S\in\mathcal{G}_2} \frac{1}{\binom{n}{|S|}} \sum_{\substack{A\in\mathcal{F}\\ A\subset S}} \frac{1}{\binom{|S|}{|A|}} \sum_{\substack{B\in\mathcal{F}\\ S\subset B}} \frac{1}{\binom{n-|S|}{n-|B|}} \leq \mathrm{E}(X) 4\sqrt{n\ln n} \frac{v-1}{\frac{n}{2} - 2\sqrt{n\ln n}}. \qquad (45)$$

Combining equality (42) with inequalities (44) and (45), we have

$$\mathrm{E}(Y) \leq \mathrm{E}(X) 4\sqrt{n \ln n} \frac{v+m-2}{\frac{n}{2} - 2\sqrt{n \ln n}}. \tag{46}$$

In particular, combining with inequality (41), we have

$$\sum_{\substack{A,B \in \mathcal{F}, |B|-|A|>1 \\ A \subset B}} \frac{|A|!(|B|-|A|)!(n-|B|)!}{n!} \leq \mathrm{E}(X) 4\sqrt{n \ln n} \frac{v+m-2}{\frac{n}{2} - 2\sqrt{n \ln n}} = o_n(\mathrm{E}(X)). \tag{47}$$

Now we consider pairs (A, B) with additional property $|B| - |A| = 1$. For any subset S, we define

$$\begin{aligned} N^+(S) &= \{T \in \mathcal{F} \mid S \subset T, |T| = |S| + 1\} \\ N^-(S) &= \{T \in \mathcal{F} \mid T \subset S, |T| = |S| - 1\}. \end{aligned}$$

Let $d^+(S) = |N^+(S)|$ and $d^-(S) = |N^-(S)|$. We have

$$\begin{aligned} \sum_{\substack{A,B \in \mathcal{F} \\ A \subset B, |B|-|A|=1}} \frac{|A|!(|B|-|A|)!(n-|B|)!}{n!} &= \sum_{\substack{A,B \in \mathcal{F} \\ A \subset B, |B|-|A|=1}} \frac{|A|!(n-|B|)!}{n!} \\ &= \sum_{A \in \mathcal{F}} \frac{d^+(A)}{\binom{n}{|A|}(n-|A|)} \qquad (48) \\ &= \sum_{B \in \mathcal{F}} \frac{d^-(B)}{\binom{n}{|B|}|B|}. \qquad (49) \end{aligned}$$

We will show most contributions to the summation above are from pairs (A, B) with $d^+(A) \geq m$ and $d^-(A) \leq v$. We define two subfamilies of $\mathcal{F}$ as follows:

$$\begin{aligned} \mathcal{F}_1 &= \{S \in \mathcal{F} \mid d^+(S) \geq m\} \\ \mathcal{F}_2 &= \{S \in \mathcal{F} \mid d^-(S) \geq v\}. \end{aligned}$$

We have

$$\begin{aligned} &\sum_{\substack{A,B \in \mathcal{F} \\ A \subset B, |B|-|A|=1}} \frac{|A|!(n-|B|)!}{n!} \\ &\leq \sum_{\substack{A \in \mathcal{F}_1, B \in \mathcal{F}_2 \\ A \subset B, |B|-|A|=1}} \frac{|A|!(n-|B|)!}{n!} + \sum_{A \in \mathcal{F} \setminus \mathcal{F}_1} \frac{d^+(A)}{\binom{n}{|A|}(n-|A|)} + \sum_{B \in \mathcal{F} \setminus \mathcal{F}_2} \frac{d^-(B)}{\binom{n}{|B|}|B|} \\ &\leq \sum_{\substack{A \in \mathcal{F}_1, B \in \mathcal{F}_2 \\ A \subset B, |B|-|A|=1}} \frac{|A|!(n-|B|)!}{n!} + \sum_{A \in \mathcal{F} \setminus \mathcal{F}_1} \frac{m-1}{\binom{n}{|A|}(n-\sqrt{2n \ln n})} \end{aligned}$$

$$+\sum_{B\in\mathcal{F}\setminus\mathcal{F}_2}\frac{v-1}{\binom{n}{|B|}(n-\sqrt{2n\ln n})}$$

$$\leq \sum_{\substack{A\in\mathcal{F}_1,B\in\mathcal{F}_2\\ A\subset B,|B|-|A|=1}}\frac{|A|!(n-|B|)!}{n!}+\frac{v+m-2}{n-\sqrt{2n\ln n}}\mathrm{E}(X). \tag{50}$$

Recall C_σ is a random full chain of subsets of $[n]$. For $i=1,2$, let $X_i=|\mathcal{F}_i\cap C_\sigma|$, so that

$$\mathrm{E}(X_i)=\sum_{F\in\mathcal{F}_i}\frac{1}{\binom{n}{|F|}}. \tag{51}$$

Since $\mathcal{F}$ is $P(G)$-free, we have $\mathcal{F}_1\cap\mathcal{F}_2=\emptyset$. In particular,

$$\mathrm{E}(X_1)+\mathrm{E}(X_2)\leq \mathrm{E}(X). \tag{52}$$

Let us consider a "diamond" configuration $S\subset A_i\subset B$ for $(i=1,2)$ with $A_1,A_2\in\mathcal{F}_1$, $B\in\mathcal{F}_2$, and $|B|=|S|+2$. In other words, $S=A_1\cap A_2$ and $B=A_1\cup A_2\in F_2$ where A_1 and A_2 $(\in\mathcal{F}_1)$ only differ by one element. For a fixed S, we define an auxiliary graph L_S with vertex set $N^+(S)\cap\mathcal{F}_1$ such that two subsets A_1,A_2 form an edge in L_S if $A_1\cup A_2\in\mathcal{F}_2$. We have

1. L_S is G-free since $\mathcal{F}$ is $P(G)$-free.
2. Each edge of L_S is in one-to-one correspondence with a diamond configuration as above.

Recall that the Turán number $t(n,G)$ is the maximum number of edges that a graph on n vertices can have without containing the subgraph G. The Erdős-Simonovits-Stone Theorem [6, 7] states

$$t(n,G)=\left(1-\frac{1}{\chi(G)-1}+o_n(1)\right)\frac{n^2}{2}. \tag{53}$$

where $\chi(G)$ is the chromatic number of G.

Let $d_1^+(S)=|N^+(S)\cap\mathcal{F}_1|$ and $d_2^-(B)=|N^-(B)\cap\mathcal{F}_2|$. The number of edges in L_S is at most $t(d_1^+(S),G)$. We have

$$\sum_S f(|S|)t(d_1^+(S),G)\geq\sum_{B\in\mathcal{F}}f(|B|-2)\binom{d_2^-(B)}{2}. \tag{54}$$

Here $f(k)$ is any nonnegative function over integers and the summation on the left is taken over all S with sizes in $(\frac{n}{2}-2\sqrt{n\ln n}-1,\frac{n}{2}+2\sqrt{n\ln n}-1)$. Choose $f(k)=\frac{1}{\binom{n}{k}(n-k)^2}$ for $k\in(\frac{n}{2}-2\sqrt{n\ln n}-1,\frac{n}{2}+2\sqrt{n\ln n}-1)$. We have

$$\sum_S f(|S|)t(d_1^+(S),G) = \left(1-\frac{1}{\chi(G)-1}+o_n(1)\right)\sum_S f(|S|)\frac{(d_1^+(S))^2}{2}$$

$$
\begin{aligned}
&\leq \frac{1}{2}\left(1-\frac{1}{\chi(G)-1}+o_n(1)\right)\sum_S f(|S|)d_1^+(S)(n-|S|)\\
&= \frac{1}{2}\left(1-\frac{1}{\chi(G)-1}+o_n(1)\right)\sum_S \frac{d_1^+(S)}{\binom{n}{|S|}(n-|S|)}\\
&= \frac{1}{2}\left(1-\frac{1}{\chi(G)-1}+o_n(1)\right)\mathrm{E}(X_1). \qquad (55)
\end{aligned}
$$

$$
\begin{aligned}
\sum_{B\in\mathcal{F}_2} f(|B|-2)\binom{d_1^-(B)}{2} &= \frac{1}{2}\sum_{B\in\mathcal{F}_2}\frac{1}{\binom{n}{|B|-2}(n-|B|+2)^2}(d_2^-(B))^2-d_2^-(B)\\
&= \frac{1}{2}\left(1+O(\frac{\sqrt{n\ln n}}{n})\right)\sum_{B\in\mathcal{F}_2}\frac{(d_2^-(B))^2-d_2^-(B)}{\binom{n}{|B|}|B|^2}\\
&= \frac{1}{2}\left(1+O(\frac{\sqrt{n\ln n}}{n})\right)\sum_{B\in\mathcal{F}_2}\frac{(d_2^-(B))^2}{\binom{n}{|B|}|B|^2}-O(\frac{1}{n})\mathrm{E}(X_2). \qquad (56)
\end{aligned}
$$

Applying the Cauchy-Schwartz Inequality. the inequalities above, and the Arithmetic-Geometric Mean Inequality, we have

$$
\begin{aligned}
\sum_{\substack{A\in\mathcal{F}_1,B\in\mathcal{F}_2\\ A\subset B,|B|-|A|=1}}\frac{|A|!(n-|B|)!}{n!} &= \sum_{B\in\mathcal{F}_2}\frac{d_2^-(B)}{\binom{n}{|B|}|B|}\\
&\leq \sqrt{\sum_{B\in\mathcal{F}_2}\frac{1}{\binom{n}{|B|}}\sum_{B\in\mathcal{F}_2}\frac{(d_B^-)^2}{\binom{n}{|B|}}|B|^2}\\
&\leq \sqrt{\mathrm{E}(X_2)\left(1-\frac{1}{\chi(G)-1}+o_n(1)\right)\mathrm{E}(X_1)}\\
&= \left(\sqrt{1-\frac{1}{\chi(G)-1}}+o_n(1)\right)\sqrt{\mathrm{E}(X_1)\mathrm{E}(X_2)}\\
&\leq \left(\sqrt{1-\frac{1}{\chi(G)-1}}+o_n(1)\right)\frac{\mathrm{E}(X_1)+\mathrm{E}(X_2)}{2}\\
&\leq \left(\sqrt{1-\frac{1}{\chi(G)-1}}+o_n(1)\right)\frac{\mathrm{E}(X)}{2}. \qquad (57)
\end{aligned}
$$

Combining inequalities (47), (50), and (57) , we have

$$
\mathrm{E}\binom{X}{2} = \sum_{\substack{A,B\in\mathcal{F},|B|-|A|>1\\ A\subset B}}\frac{|A|!(|B|-|A|)!(n-|B|)!}{n!}
$$

$$+ \sum_{\substack{A,B\in\mathcal{F},|B|-|A|=1\\ A\subset B}} \frac{|A|!(|B|-|A|)!(n-|B|)!}{n!}$$

$$\begin{aligned} &\leq o_n(\mathrm{E}(X)) + \left(\sqrt{1-\frac{1}{\chi(G)-1}} + o_n(1)\right)\frac{\mathrm{E}(X)}{2} \\ &\leq \left(\sqrt{1-\frac{1}{\chi(G)-1}} + o_n(1)\right)\frac{1}{2}\mathrm{E}(X). \end{aligned} \tag{58}$$

Combining inequalities (39) and (58), we have

$$E(X) \leq 1 + \sqrt{1-\frac{1}{\chi(G)-1}} + o_n(1). \tag{59}$$

The proof is finished by observing $|\mathcal{F}| \leq \mathrm{E}(X)\binom{n}{\lfloor \frac{n}{2}\rfloor}$. $\square$

4 Further research

Let

$$\pi(H) := \lim_{n\to\infty} \frac{\mathrm{La}(n,H)}{\binom{n}{\lfloor \frac{n}{2}\rfloor}},$$

when this limit exists. Does this limit exist for all posets H, and, if so, how does it depend on H? For posets H of height two, we know that the limit, when it exists, belongs to the interval $[1,2]$. Are there any H of height two such that $\pi(H)$ is strictly between 1 and 2?

More generally, for all posets H where we know $\pi(H)$, $\pi(H)$ is an integer. Is this true in general? In fact, examples we looked have have $\pi(H)$ equal to the maximum number m such that the middle m levels of the Boolean lattice $B_n = (2^{[n]}, \subseteq)$ do not contain H, no matter how large n is (as observed by Mike Saks and Pete Winkler, unpublished).

We once asked whether there exists a number c_h such that for all posets H of height h, $\pi(H) \leq c_h$. As we noted above, $c_2 = 2$. However, Lu and, independently, Tao Jiang, pointed out that no such c_h for $h \geq 3$. The idea is that if one takes $\mathcal{F}$ to consist of the middle m levels in the Boolean lattice B_n, then two sets $A, B \in \mathcal{F}$ with $A \subset B$ have at most $2^{m-1} - 2$ sets C with $A \subset C \subset B$. Hence, the family $\mathcal{F}$, which has size $\sim m\binom{n}{\lfloor \frac{n}{2}\rfloor}$, avoids the height 3 poset consisting of a minimum element, a maximum element, and an antichain of $2^{m-1}-1$ elements in between. This forces c_3 to be larger than any m, so that no such c_3 exits. It seems that not just the height, but the width, of H affects $\pi(H)$.

It would therefore be interesting to determine $\pi(B_n)$ for the Boolean lattice B_n . The smallest crown for which π is not yet determined is $\mathcal{O}_6$, the height two poset formed by the middle two levels of B_3. Even for a poset as fundamental as the diamond poset B_2, we only know that $\pi(B_2)$, if it exists, must be in the interval $[2,3]$.

References

[1] B. Bukh, Set families with a forbidden poset, preprint (2008).

[2] H. Chernoff, A note on an inequality involving the normal distribution, *Ann. Probab.* **9** (1981), 533-535.

[3] A. De Bonis, G. O.H. Katona, Largest families without an r-fork, *Order* **24** (2007), 181–191.

[4] A. De Bonis, G. O.H. Katona, K. J. Swanepoel, Largest family without $A \cup B \subset C \cap D$, *J. Combin. Theory (Ser. A)* **111** (2005), 331-336.

[5] P. Erdős, On a lemma of Littlewood and Offord, *Bull. Amer. Math. Soc.* **51** (1945), 898-902.

[6] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 51-57.

[7] P. Erdős and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087-1091.

[8] C. M. Fortuin, P. N. Kasteleyn, and J. Ginibre, Correlation inequalities for some partially ordered sets, *Comm. Math. Physics*, **22** (1971), 89-103.

[9] J. R. Griggs, G. O. H. Katona, No four subsets forming an N, *J. Combinatorial Theory (Ser. A)* **115** (2008), 677–685.

[10] E. Sperner, Ein Satz über Untermegen einer endlichen Menge, *Math. Z.* **27** (1928), 544-548.

[11] H. T. Thanh, An extremal problem with excluded subposets in the Boolean lattice, *Order* **15** (1998), 51-57.