Forbidden subposet problems

Balázs Patkós

A family $\mathcal{G}$ of sets is a *weak copy* of a poset $(P, \leqslant)$, if there exists a bijection $f : P \to \mathcal{G}$ such that $p \leqslant q$ implies $f(p) \subseteq f(q)$, and $\mathcal{G}$ is a strong copy of $(P, \leqslant)$, if there exists a bijection $f : P \to \mathcal{G}$ such that $p \leqslant q$ if and only if $f(p) \subseteq f(q)$. With some abuse of notation, we will write P instead of $(P, \leqslant)$ and say that a family $\mathcal{F}$ is *weak (strong) P-free* if it does not contain weak (strong) copies of P. The *forbidden subposet problem* of determining $\mathrm{La}(n, P)$ $(\mathrm{La}^*(n, P))$, the maximum size of a weak (strong) P-free family $\mathcal{F} \subseteq 2^{[n]} := \{S : S \subseteq \{1, 2 \ldots, n\}\}$. was introduced by Katona and Tarján [8] in the early 80s.

Let $e(P)$ $[e^*(P)]$ denote the largest integer k such that for any n, the middle k layers of $\mathcal{B}_n$, is [induced] P-free. By definition, we have

$$\Sigma\big(n, e(P)\big) \leq \mathrm{La}(n, P) \qquad \text{and} \qquad \Sigma\big(n, e^*(P)\big) \leq \mathrm{La}^{**}(n, P).$$

Griggs and Lu [7] and independently Bukh [3] conjectured in 2009 that for any poset P, $\mathrm{La}(n, P) = \big(e(P) + o(1)\big)\binom{n}{\lfloor n/2 \rfloor}$ and proved this conjecture for important classes of posets.

In order to describe Bukh's result, we define a *tree poset* T to be one such that its unoriented Hasse diagram is a tree. The *height* $h(P)$ of a poset P is the length of the longest chain in P. Observe that for any tree poset, it is the case that $e(P) = h(P) - 1$.

Theorem 0.1 (Bukh [3]). *For any tree poset T, we have* $\mathrm{La}(n, T) = \big(h(T) - 1 + O_T\big(\frac{1}{n}\big)\big)\binom{n}{\lfloor n/2 \rfloor}$.

The same result for the asymptotics of $\mathrm{La}^*(n, T)$ is by Boehnlein and Jiang [1] but their error term is $o_T(1)$ instead of $O_T\big(\frac{1}{n}\big)$.

Theorem 0.2 (LYM inequality [9, 13, 11]). *If $\mathcal{F} \subseteq 2^{[n]}$ is a k-Sperner (i.e. $\mathcal{C}_{k+1}$-free) family, then*

$$\lambda_n(\mathcal{F}) \overset{\text{def}}{=} \sum_{F \in \mathcal{F}} \binom{n}{|F|}^{-1} \leq k.$$

The quantity $\lambda_n(\mathcal{F})$ is called the *Lubell mass* of $\mathcal{F}$ is one of the major and most common used tools for establishing upper bounds on $\mathrm{La}(n, P)$. The connection between upper bounds on $\lambda_n(\mathcal{F})$ and the size of $\mathcal{F}$ is made in the following simple lemma.

Lemma 1. *If for some* $\mathcal{F} \subseteq 2^{[n]}$ *and positive real* c*, we have* $\lambda_n(\mathcal{F}) \le c$*, then* $|\mathcal{F}| \le c\binom{n}{\lfloor n/2 \rfloor}$*. Moreover, if* $c = k$ *is an integer, then* $|\mathcal{F}| \le \Sigma(n,k)$*.*

As every P-free family is $\mathcal{C}_{|P|}$-free, another immediate corollary of Theorem 0.2 is that for any P-free family $\mathcal{F} \subseteq 2^{[n]}$, we have $\lambda_n(\mathcal{F}) \le |P| - 1$. An induced analog of this inequality resisted the attempts of researchers for long time, until Méroueh [10] proved the following.

Theorem 0.3 (Méroueh [10]). *For an* P *there exists a constant* $C = C(P)$ *such that for any induced-*P*-free family* $\mathcal{F} \subseteq 2^{[n]}$ *we have* $\lambda_n(\mathcal{F}) \le C$*.*

Note that Theorem 0.3, via Lemma 1, establishes Methuku and Pálvölgyi's theorem [12] that $\mathrm{La}^*(n,P) = O\big(\binom{n}{\lfloor n/2 \rfloor}\big)$.

In probabilistic terminology, the Lubell mass is the expected value $\mathbb{E}[X]$ of the random variable $X = X_{\mathcal{F}}$ which counts the number of sets $F \in \mathcal{F}$ that a maximal chain picked uniformly at random contains. Using a bound on the average number of sets in $\mathcal{F} \cap \mathcal{C}$ was addressed from early on. However, the method that is now known as the *chain partitioning method* was introduced by Griggs, Li, and Lu [6] and a detailed discussion appeared in the paper by Griggs and Li [5].

The method can be very simply described in general. Fix a family $\mathcal{F} \subseteq 2^{[n]}$. Let $\mathbf{C}$ denote the set of all maximal chains in $[n]$ and let $\mathbf{C}^1 \cup \mathbf{C}^2 \cup \dots \cup \mathbf{C}^m$ be a partition of $\mathbf{C}$. Then one can introduce the corresponding random variables $X^1, X^2, \dots, X^m$ with X^j being the number of sets $F \in \mathcal{F}$ that a maximal chain picked uniformly at random from $\mathbf{C}^j$ contains. Then clearly, the Lubell mass $\lambda_n(\mathcal{F})$ is the weighted average $\sum_{j=1}^m \frac{|\mathbf{C}^j|}{n!}\mathbb{E}(X^j)$. In particular, if $\mathbb{E}[X^j] \le C$ holds for $1 \le j \le m$, then $\lambda_n(\mathcal{F}) \le C$.

Commonly used partitions are

- the *min-partition* $\mathbf{C}_\emptyset \cup \bigcup_{F \in \mathcal{F}} \mathbf{C}_F^{\min}$, where $\mathbf{C}_F^{\min}$ is the set of those maximal chains C for which F is the smallest set in $\mathcal{F}$ that lies in C, and $\mathbf{C}_\emptyset = \{\mathcal{C} \in \mathbf{C} : \mathcal{C} \cap \mathcal{F} = \emptyset\}$,
- the *max-partition* $\mathbf{C}_\emptyset \cup \bigcup_{F \in \mathcal{F}} \mathbf{C}_F^{\max}$, where $\mathbf{C}_F^{\max}$ is the set of those maximal chains C for which F is the largest set in $\mathcal{F}$ that lies in C,
- the *min-max-partition* $\mathbf{C}_\emptyset \cup \bigcup_{F,F' \in \mathcal{F}, F \subseteq F'} \mathbf{C}_{F,F'}$, where $\mathbf{C}_{F,F'}$ is the set of those maximal chains C for which F is the smallest and F' is the largest element of $\mathcal{F}$ that each lie in C, in particular, if $F = F'$, then F is the only set in $\mathcal{F} \cap \mathcal{C}$.

An application.

Theorem 0.4. $\mathrm{La}(n, \diamond_3) = \Sigma(n,3)$*.*

Proof. Consider the min-max partition of a $\diamond_3$-free family $\mathcal{F}$. By the $\diamond_3$-free property, for any $F \subset F'$ with $F, F' \in \mathcal{F}$, there are at most 2 other sets $G \in \mathcal{F}$ with $F \subset G \subset F'$. So the average number of sets in a chain in $\mathbf{C}_{F,F'}$ is at most $2 + \frac{2}{|F'|-|F|} \le 3$. Therefore, $\lambda_n(\mathcal{F}) \le 3$ and Lemma 1 finishes the proof.

Now we consider the construction of Ellis, Ivan, and Leader [4] that serves as a counterexample to the conjectures that $\mathrm{La}(n,P) = \big(e(P)+o(1)\big)\binom{n}{\lfloor n/2\rfloor}$ and $\mathrm{La}^*(n,P) = \big(e^*(P)+o(1)\big)\binom{n}{\lfloor n/2\rfloor}$.

In fact, this construction disproved the so-called *Daisy Conjecture* of Bollobás, Leader, and Malvenuto [2] and, in fact, disproved many (seemingly unrelated) conjectures in extremal set theory.

To formulate the statement, let us define an r-*daisy* to be the family of all six r-sets that contain some fixed $(r-2)$-set R (the *stem* of the daisy) and that are subsets of a fixed $(r+2)$-set $T \supset R$. Let $\mathrm{ex}(n, D_r)$ be the maximum number of r-sets in a family $\mathcal{F}$ of r-sets on vertex set $[n]$ such that $\mathcal{F}$ does not contain any r-daisies and let $\pi_r = \lim_{n\to\infty} \frac{\mathrm{ex}(n,D_r)}{\binom{n}{r}}$. The Daisy conjecture stated that π_r tends to 0 as r tends to infinity. Bollobás, Leader, and Malvenuto claimed that the Daisy conjecture is equivalent to the statement that $\frac{\mathrm{ex}(2n,D_n)}{\binom{2n}{n}}$ tends to 0 as n tends to infinity.

Lemma 2 (Bollobás, Leader, Malvenuto [2]). *For any $c \geq 0$,*

$$\liminf_{r\to\infty} \pi_r > c \qquad \Longleftrightarrow \qquad \liminf_{n\to\infty} \frac{\mathrm{ex}(2n, D_n)}{\binom{2n}{n}} > c.$$

Proof. Assume first that $\liminf \pi_r$ tends to c, so we can pick an r with $\pi_r < c + \varepsilon/2$ and thus for large enough n, we have $\frac{\mathrm{ex}(n,D_r)}{\binom{n}{r}} < c+\varepsilon$. So if $\mathcal{F} \subseteq \binom{[2n]}{n}$ does not contain an n-daisy, then we count pairs (G,F) with $G \subseteq F \in \mathcal{F}$ and $|G| = n-r$. Fixing F, we get that the number of such pairs is exactly $|\mathcal{F}|\binom{n}{n-r}$. Fixing G, the r-uniform family $\mathcal{F}_G := \{F \setminus G \in \mathcal{F} : G \subseteq F\}$ lives on $n+r$ vertices and does not contain any r-daisy, so $|\mathcal{F}| \leq \mathrm{ex}(n+r, D_r) < (c+\varepsilon)\binom{n+r}{r}$. Therefore the number of pairs is at most $\binom{2n}{n-r}(c+\varepsilon)\binom{n+r}{r}$. Therefore,

$$|\mathcal{F}|\binom{n}{n-r} < \binom{2n}{n-r}(c+\varepsilon)\binom{n+r}{r}$$

and so

$$|\mathcal{F}| < (c+\varepsilon)\frac{\binom{2n}{n-r}\binom{n+r}{r}}{\binom{n}{n-r}} = (c+\varepsilon)\binom{2n}{n}$$

as needed.

Assume next $\liminf \frac{\mathrm{ex}(2n,D_n)}{\binom{2n}{n}}$ tends to c. As a result, for any $\varepsilon > 0$ and large enough r, we have $\frac{\mathrm{ex}(2r,D_r)}{\binom{2r}{r}} < c + \varepsilon$. So for large enough n and $\mathcal{F} \subseteq \binom{[n]}{r}$ with no r-daisy, we count pairs (F,G) with $\mathcal{F} \ni F \subseteq G$ and $|G| = 2r$. Fixing F first yields that the number of such pairs is $|\mathcal{F}|\binom{n-r}{r}$. Fixing G, we know that $\mathcal{F}^G := \{F \in \mathcal{F} : F \subset G\}$ contains no r-daisies, so $|\mathcal{F}^G| \le \mathrm{ex}(2r, D_r) \le (c+\varepsilon)\binom{2r}{r}$. So the number of pairs is at most $\binom{n}{2r}(c+\varepsilon)\binom{2r}{r}$. Therefore,

$$|\mathcal{F}|\binom{n-r}{r} \le \binom{n}{2r}(c+\varepsilon)\binom{2r}{r}$$

and so

$$|\mathcal{F}| \le (c+\varepsilon)\frac{\binom{n}{2r}\binom{2r}{r}}{\binom{n-r}{r}} = (c+\varepsilon)\binom{n}{r}$$

as needed.

The construction by Ellis, Ivan and Leader [4] establishes that there is a family in the middle layer of the Boolean lattice which has nontrivial density.

Theorem 0.5 (Ellis, Ivan, and Leader [4])**.**

$$\liminf_{r\to\infty} \pi_r \ge \prod_{k=1}^{\infty}\left(1-2^{-k}\right) \ge 0.288787.$$

Consequently, $\big(0.288787 - o(1)\big)\binom{2n}{n} \le \mathrm{ex}(2n, D_n)$.

Proof. Here we only sketch the main points of the proof, leaving the details to the reader. First, define $\mathcal{F}_{r,2^r-1}$ as follows: fix a bijection f between the nonzero vectors of the vector field $\mathbb{F}_2^r$ and $[2^r-1]$, and let

$$\mathcal{F}_{r,2^r-1} = \big\{f(B) : B \text{ is a basis of } \mathbb{F}_2^r\big\}.$$

By counting the number of ordered bases of $\mathbb{F}_2^r$, one can obtain that $\left|\mathcal{F}_{r,2^r-1}\right|/\binom{2^r-1}{r} > \prod_{k=1}^{r}\left(1-2^{-k}\right)$.

The next goal is to show that any r-daisy in $\left[2^r-1\right]$ contains an r-set corresponding to a linearly dependent set of r vectors. Suppose for a contradiction that there exists some r-daisy F in $\left[2^r-1\right]$ whose r-sets all correspond to linearly independent sets of r-vectors. Thus, the stem of the daisy forms the basis of an $(r-2)$-dimensional subspace W. Let $\{u_1, u_2, u_3, u_4\}$ be the remaining vertices of the daisy. Modding out by W, $\mathbb{F}_2^r/W$ is a 2-dimensional subspace over $\mathbb{F}_2$. Since there is no $u_i \in W$, it must be the case that there are distinct u_i, u_j with $f^{-1}\big(u_i\big) + W = f^{-1}\big(u_j\big) + W$, contradicting the fact that $f^{-1}\big(u_i\big)$, $f^{-1}\big(u_j\big)$, and the vectors corresponding to the vertices of the stem form an independent set.

Next, let $n \ge 2^r - 1$ and the family $\mathcal{F}_{r,n}$ be the blow-up of $\mathcal{F}_{r,2^r-1}$. That is, we partition $[n]$ into 2^r-1 classes $X_1, X_2, \ldots, X_{2^r-1}$ as equally as possible. The index $i(x)$ of $x \in [n]$ is defined by $x \in X_{i(x)}$. Then let

$$\mathcal{F}_{r,n} = \Big\{\big\{x_1, x_2, \ldots, x_r\big\} : \big\{i(x_1), i(x_2), \ldots, i(x_r)\big\} \in \mathcal{F}_{r,2^r-1}\Big\}.$$

This is also r-daisy-free and the lower bound on $\lim\limits_{r\to\infty} \pi_r$ now follows.

References

[1] Boehnlein, E., and Jiang, T. Set families with a forbidden induced subposet. *Combinatorics, Probability and Computing 21*, 04 (2012), 496–511.

[2] Bollobás, B., Leader, I., and Malvenuto, C. Daisies and other Turán problems. *Combin. Probab. Comput. 20*, 5 (2011), 743–747.

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

[4] Ellis, D., Ivan, M.-R., and Leader, I. Turán densities for daisies and hypercubes. *Bull. Lond. Math. Soc. 56*, 12 (2024), 3838–3853.

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

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

[7] Griggs, J. R., and Lu, L. On families of subsets with a forbidden subposet. *Combinatorics, Probability and Computing 18*, 05 (2009), 731–748.

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

[9] Lubell, D. A short proof of Sperner's lemma. *Journal of Combinatorial Theory 1*, 2 (1966), 299.

[10] Méroueh, A. A LYM inequality for induced posets. *Journal of Combinatorial Theory, Series A 155* (2018), 398–417.

[11] Meshalkin, L. D. Generalization of Sperner's theorem on the number of subsets of a finite set. *Theory of Probability & Its Applications 8*, 2 (1963), 203–204.

[12] Methuku, A., and Pálvölgyi, D. Forbidden hypermatrices imply general bounds on induced forbidden subposet problems. *Combinatorics, Probability and Computing* (2017), 1–10.

[13] Yamamoto, K. Logarithmic order of free distributive lattice. *Journal of the Mathematical Society of Japan 6*, 3-4 (1954), 343–353.