Progress on Poset-free Families of Subsets

Jerrold R. Griggs * Wei-Tian Li †

October 1, 2017

Abstract

Increasing attention is being paid to the study of families of subsets of an n-set that contain no subposet P. Especially, we are interested in such families of maximum size given P and n. For certain P this problem is solved for general n, while for other P it is extremely challenging to find even an approximate solution for large n. It is conjectured that for any P, the maximum size is asymptotic to a constant times $\binom{n}{\lfloor \frac{n}{2} \rfloor}$, where the constant is a certain integer depending on P. This survey has two purposes. First, we want to bring this exciting line of research to the attention of a wider audience. Second, we want to make experts aware of the broad range of recent progress in the area.

1 First Results: Families Without Chains

Problems of forbidding given structures in a larger structure are very popular in extremal combinatorics. The study of forbidden subgraph problems, starting from Mantel's result on triangle-free graphs, is nowadays a well-developed discipline. The Turán theory on hypergraphs contains beautiful theorems and challenging unsolved problems.

A research area that has become fertile in recent years considers forbidding poset structures in families of subsets. Let $P = (P, \leq)$ be a finite poset. We say that another poset Q contains P as a *subposet* (in the weak sense), if there exists an order-preserving injection i from P to Q. We view a family $\mathcal{F}$ of subsets of a finite set as a poset itself, ordered according to the inclusion relation of sets. More precisely, working in the Boolean lattice $\mathcal{B}_n$ of all 2^n subsets of $[n] := \{1, 2, \cdots, n\}$, ordered by inclusion, we want to understand how large a family of subsets in $\mathcal{B}_n$ can be without containing a given poset P as a subposet. We let $\binom{[n]}{k}$ denote the collection of k-subsets of $[n]$.

The foundational result in the area dates back to 1928. Sperner studied *antichains*, which are families of sets without any set containing another, and obtained the following:

*Department of Mathematics, University of South Carolina, Columbia, SC, USA 29208 (griggs@math.sc.edu). Research supported in part by a grant from the Simons Foundation (#282896 to Jerrold Griggs) and by a long-term visiting position at the IMA, University of Minnesota.

†Department of Applied Mathematics, National Chung Hsing University, Taichung, Taiwan 40227 (weitianli@dragon.nchu.edu.tw). Supported by Ministry of Science and Technology (No. 103-2115-M-005 -003 -MY2)

Theorem 1.1 [49] *Let $\mathcal{F}$ be an antichain of subsets of $[n]$. Then $|\mathcal{F}|$ is at most the largest binomial coefficient, namely $\binom{n}{\lfloor \frac{n}{2} \rfloor}$. Moreover, the only antichains achieving the bound are $\binom{[n]}{\lfloor \frac{n}{2} \rfloor}$ and $\binom{[n]}{\lceil \frac{n}{2} \rceil}$.*

Note that we get two different extremal families, or just one (the two listed are identical), depending on the parity of n.

A generalization of Sperner's theorem was given by Erdős, who developed it in the 1940's to make progress on a question of Littlewood and Offord concerning the distribution of roots of random polynomials. The *size* of a poset is its cardinality. A poset is a *chain* if any two of its elements are comparable. Erdős gave the following result on forbidding chains of a given size: .

Theorem 1.2 [18] *Let $k \geq 1$. Let $\mathcal{F}$ be a family of subsets of $[n]$ such that no $k+1$ subsets form a chain. Then $|\mathcal{F}|$ is at most the sum of the k middle binomial coefficients in n. Moreover, the only families achieving the bound are obtained by taking all subsets of the k middle sizes in $[0, n]$.*

Again, the maximum size of such a family $\mathcal{F}$ is given for all n, and there is just one, or two, extremal families. Since these bounds and the families of Erdős's Theorem occur repeatedly in the theory, it is useful to adopt notation from [23]: Let $\mathcal{B}(n, k)$ be a family of subsets of $[n]$ of the k middle sizes, $\binom{[n]}{\lfloor (n-k+1)/2 \rfloor} \cup \cdots \cup \binom{[n]}{\lfloor (n+k-1)/2 \rfloor}$ or $\binom{[n]}{\lceil (n-k+1)/2 \rceil} \cup \cdots \cup \binom{[n]}{\lceil (n+k-1)/2 \rceil}$. So $\mathcal{B}(n, k)$ is one or two possible families, depending on the parity of $n + k$. Also, let $\Sigma(n, k)$ denote $|\mathcal{B}(n, k)|$.

By saying no $k + 1$ elements in a family $\mathcal{F}$ form a chain, it means that $\mathcal{F}$ does not contain any $\mathcal{P}_{k+1}$, where $\mathcal{P}_r$ denotes the *chain poset* (also called *path poset*) of size r.

In the early 1980s, Katona started to investigate forbidding posets other than chains in the Boolean lattice. Since then a rapidly growing number of results on such problems have been discovered by Katona, his collaborators, and other researchers who are inspired by the pioneering work. Here we attempt to collect this work, particularly the research papers published in past ten years. We seek to identify the core ideas and prospective directions for continuing study. Interested readers should also see other surveys in recent years, especially [28, 23, 36, 30].

2 Forbidding Other Posets

Following Katona we denote by $\mathrm{La}(n, P)$ the maximum size of a family $\mathcal{F}$ of subsets of $[n]$ containing no subposet P. If a family of subsets does not contain P as a subposet, then we say it is *P-free*. Note that we are forbidding P as a subposet in the weak sense, meaning that we are excluding not only P itself, but also any poset containing P. For instance, let $\mathcal{V}_r$ denote the *fork poset* with one element below the other r elements. When we forbid $\mathcal{V}_r$, we are forbidding any subset being in r others in the family, whether or not those others contain each other. We are interested in the exact value of $\mathrm{La}(n, P)$, when we can determine it, or else we are happy if we can determine its leading asymptotic behavior, for fixed poset P, as $n \to \infty$.

One poset parameter we sometimes need is the *height* of poset P, denoted $h(P)$, which is the largest size of any chain in P.

The first small poset P that was investigated was $\mathcal{V} = \mathcal{V}_2$ (which has a Hasse diagram that looks like the letter V). Katona and Tarján (1981) gave the asymptotic behavior of $\mathrm{La}(n, \mathcal{V})$:

Theorem 2.1 [32] *As $n \to \infty$,*

$$1 + \frac{1}{n} + \Omega\left(\frac{1}{n^2}\right) \leq \frac{\mathrm{La}(n, \mathcal{V})}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq 1 + \frac{2}{n}.$$

Note that $\mathcal{B}(n, 1)$, the subsets of middle size in n, is a $\mathcal{V}$-free family of asymptotically optimal size. However, optimal families are slightly larger (and not known)–we can add some sets in the level above the middle one while remaining $\mathcal{V}$-free. The lower bound in the theorem adds in a lower order term, based on a construction of Graham and Sloane [19] in coding theory. The error term in the upper bound is twice that of the lower bound. It would be a significant accomplishment to improve the error term and determine constant c such that $\mathrm{La}(n, \mathcal{V}) = (1 + \frac{c}{n} + o(\frac{1}{n}))\binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Thanh (1998) generalized the result of Katona and Tarján on poset $\mathcal{V}$ to the r-fork, and Katona and DeBonis later improved the error term:

Theorem 2.2 [50, 13] *For general r, as $n \to \infty$,*

$$1 + \frac{r-1}{n} + \Omega\left(\frac{1}{n^2}\right) \leq \frac{\mathrm{La}(n, \mathcal{V}_r)}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq 1 + 2\frac{r-1}{n} + O\left(\frac{1}{n^2}\right).$$

Another way to build up the $\mathcal{V}$ poset is to add a second element below the top two. Called the *butterfly poset* $\bowtie$, it has two minimal elements, each below each of two maximal elements. We have a jump in the possible size of P-free families for this poset compared to $\mathcal{V}$. Particularly surprising is that for the butterfly poset, the value of $\mathrm{La}(n, P)$ can be determined precisely. Observe that for any k, two distinct k-sets have at most one $(k-1)$-subset below them both. It means that the union of two consecutive levels in $\mathcal{B}_n$ is $\bowtie$-free. DeBonis, Katona, and Swanepoel (2005) used the cyclic permutation method to prove that the family $\mathcal{B}(n, 2)$ is as large as possible, provided $n \geq 3$:

Theorem 2.3 [14] *For butterfly-free families, when $n \geq 3$,*

$$\mathrm{La}(n, \bowtie) = \Sigma(n, 2).$$

For $n \geq 5$, the only extremal families are $\mathcal{B}(n, 2)$.

Omitting one ordered pair in the butterfly poset gives a poset that is intermediate between $\mathcal{V}$ and $\bowtie$. Called $\mathcal{N}$ because of its Hasse diagram, it has elements corresponding to subsets a, b, c, d with $a \subset b, c \subset b, c \subset d$. It immediately satisfies $\mathrm{La}(n, \mathcal{V}) \leq \mathrm{La}(n, \mathcal{N}) \leq \mathrm{La}(n, \bowtie)$. Griggs and Katona (2008) discovered for this poset that $\mathrm{La}(n, P)$ grows like $\mathrm{La}(n, \mathcal{V})$:

Theorem 2.4 [20] *For $\mathcal{N}$-free families, as $n \to \infty$,*

$$1 + \frac{1}{n} + \Omega\left(\frac{1}{n^2}\right) \leq \frac{\mathrm{La}(n, \mathcal{N})}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq 1 + \frac{2}{n} + O\left(\frac{1}{n^2}\right).$$

What if we take the butterfly poset and add more elements to the two levels? Say we have r elements each above each of s elements. Call this the *complete 2-level poset* $\mathcal{K}_{r,s}$, named after the graph that looks like its Hasse diagram. For $r = s = 2$ it is the butterfly. DeBonis and Katona found bounds on $\mathrm{La}(n, K_{r,s})$ that are asymptotically similar to what was found for the butterfly.

Theorem 2.5 [13] *Let $r \geq s \geq 2, r \geq 3$. For $K_{r,s}$-free families, as $n \to \infty$, we have*

$$\Sigma(n, 2) + \left(2\ \frac{r+s-4}{n} + \Omega\left(\frac{1}{n^2}\right)\right)\binom{n}{\lfloor \frac{n}{2} \rfloor}$$
$$\leq\ \mathrm{La}(n, K_{r,s}) \leq \left(2 + 2\ \frac{r+s-3}{n} + O\left(\frac{1}{n^2}\right)\right)\binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

Their upper bound for $\mathrm{La}(\mathcal{K}_{r,s})$ follows from a method of partitioning a $\mathcal{K}_{r,s}$-free families into the union of a $\mathcal{V}_r$-free family and a Λ_s–free family, where the *s-lambda poset* Λ_s is the dual of $\mathcal{V}_s$.

Let us mention that recently Patkós [46] has obtained results for the complete 3-level poset $\mathcal{K}_{r,s,t}$, where the levels have sizes r, s, t, respectively.

Griggs and Lu [24] considered several other families of posets, including batons and height two trees. They also considered a different generalization of the butterfly poset. The *crown poset* $\mathcal{O}_{2k}$, $k \geq 2$, has Hasse diagram that is an up-and-down cycle on $2k$ vertices, namely, $a_1 \leq b_1 \geq a_2 \leq b_2 \cdots \geq a_k \leq b_k \geq a_1$. The smallest crown is $\mathcal{O}_4$, which is the butterfly, while $\mathcal{O}_6$ is the middle two levels of the Boolean lattice $\mathcal{B}_3$. Their 2008 work also introduced more probabilistic methods into the subject. They managed to determine $\mathrm{La}(n, \mathcal{O}_{2k})$ asymptotically for all even $k \geq 2$. Later on (2014), Lu managed to do this as well for odd $k \geq 7$.

Theorem 2.6 [24, 37] *Let $k = 4$ or ≥ 6. For $\mathcal{O}_{2k}$-free families,*

$$1 + \frac{1}{n} + \Omega\left(\frac{1}{n^2}\right) \leq \frac{\mathrm{La}(n, \mathcal{O}_{2k})}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq 1 + \frac{2}{n} + O\left(\frac{1}{n^2}\right).$$

It remains open to determine $\mathrm{La}(n, P)$ asymptotically for $\mathcal{O}_6$ and $\mathcal{O}_{10}$. The best upper bound is larger than above by a factor of $(1 + \frac{1}{\sqrt{2}})$. In contrast to the butterfly ($k = 2$), we see that for all $k \geq 6$, $\mathrm{La}(n, \mathcal{O}_{2k}) \sim \binom{n}{\lfloor \frac{n}{2} \rfloor}$. This holds as well for $k = 4$, and likely is true for $k = 3, 5$.

The next section describes the difficulty determining $\mathrm{La}(n, P)$ when P is a four-element diamond poset. Li (2011) then weakened the diamond slightly, considering the poset on elements a, b, c, d with $a < b, a < c < d$, denoted by $\mathcal{J}$ (J) due to its Hasse diagram. Forbidding $\mathcal{J}$ is weaker than forbidding the chain $\mathcal{P}_3$, but stronger than forbidding $\mathcal{P}_4$ or the four-element diamond. It was surprising that there is a nice answer here.

Theorem 2.7 [36] *For $\mathcal{J}$-free families,*

$$\mathrm{La}(n, \mathcal{J}) = \Sigma(n, 2).$$

All posets P mentioned so far are *ranked*, which means that there is a function r from P to the nonnegative integers such that whenever element y covers x, $r(y) = r(x) + 1$. For instance, the Boolean lattice $\mathcal{B}_n$ is ranked by taking $r(X) = |X|$ for a subset X of $[n]$. The smallest unranked posets contain five elements. Methuku and Tompkins considered one such poset, with partial order relations: $a_1 < a_2,\ b_1 < b_2 < b_3,\ a_1 < b_3,\ b_1 < a_2$. They named it the *skew-butterfly* and determined $\mathrm{La}(n, P)$ for this poset:

Theorem 2.8 [44] *Let P be the skew-butterfly. For $n \geq 3$,* $\mathrm{La}(n, P) = \Sigma(n, 2)$.

3 Searching for Diamonds

After speaking about the early work on forbidden subposets back in March, 2007, including the butterfly and $\mathcal{N}$ posets, Griggs was asked about another poset with a 4-cycle Hasse diagram, the Boolean lattice $\mathcal{B}_2$. Because of its shape, it makes sense to call it the *diamond poset*. More generally, let $\mathcal{D}_k$ denote the poset with one maximum element above the rest, one minimum element below the rest, and k elements in between. We call this the *k-diamond poset*, so that $\mathcal{D}_2$ denotes the diamond poset. People asked Katona about $\mathcal{D}_2$ at talks before that. He wrote [30] that when he was asked in 2004 about $\mathrm{La}(n, \mathcal{D}_2)$, he answered, "*This is a good question, I think the present method will work on this problem, too.*" To the contrary, rather unexpectedly, it turns out that solving $\mathrm{La}(n, \mathcal{D}_2)$ has become one of the most challenging problems in extremal combinatorics.

Immediately, because $\mathcal{D}_2$ contains a 3-chain, $\mathcal{B}(n, 2)$ is diamond-free. On the other hand, since $\mathcal{D}_2$ is a subposet of a 4-chain, by Erdős's Theorem 1.2, no diamond-free family is larger than $\mathcal{B}(n, 3)$. It means that $\Sigma(n, 2) \leq \mathrm{La}(n, \mathcal{D}_2) \leq \Sigma(n, 3)$, so asymptotically

$$2(1 - o(1)) \leq \frac{\mathrm{La}(n, \mathcal{D}_2)}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq 3.$$

An easy averaging argument reduces the upper bound to 2.5. Griggs and Lu believed that upper bound to be soft– the lower bound of 2 should be the answer asymptotically. While still a student, Li joined the effort to improve the upper bound.

Using averaging arguments and the Lubell function (which we introduce later), Griggs, Li, and Lu brought the upper bound strictly below 2.3. They announced their progress at conferences, and the group of Axenovich, Manske, and Martin responded by reducing the bound. Griggs *et al.* then did more computations to push their argument for an even better bound. Later Kramer, Martin, and Young succeeded in reducing it (asymptotically) to 2.25, which is the barrier for methods depending on the bounding the Lubell function of diamond-free families. It is then important that more recently the bound was brought down even lower, to 2.208, by Grósz, Methuku, and Tompkins. They used a partitioning of the set of maximal chains, and they introduced an induction method. Theirs is the current best bound, still well above the value 2 that most researchers expect is the actual value. We summarize these results here:

Theorem 3.1 *For $\mathcal{D}_2$-free families, we have for sufficiently large n*

(1)[23] $\mathrm{La}(n, \mathcal{D}_2) \leq 2.296\binom{n}{\lfloor \frac{n}{2} \rfloor}$,

(2)[2] $\mathrm{La}(n, \mathcal{D}_2) \leq 2.283\binom{n}{\lfloor \frac{n}{2} \rfloor}$,

(3)[23] $\mathrm{La}(n, \mathcal{D}_2) \leq 2.273\binom{n}{\lfloor \frac{n}{2} \rfloor}$,

(4)[35] $\mathrm{La}(n, \mathcal{D}_2) \leq 2.25(1 + o(1))\binom{n}{\lfloor \frac{n}{2} \rfloor}$.

(5)[27] $\mathrm{La}(n, \mathcal{D}_2) \leq 2.208\binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Since this problem of determining the largest size of a diamond-free family $\mathcal{F}$ of subsets of $[n]$ is so daunting, how about restricting $\mathcal{F}$ and seeing if the bound of 2 applies? Various researchers independently considered $\mathcal{F}$ restricted to just the three middle sizes, that is, $\mathcal{F} \subseteq \mathcal{B}(n, 3)$. A diamond in such a family consists of a subset A in the lowest level, a subset C in the highest level, and both subsets B obtained by adding an element of $C - A$ to A. Moreover, this problem appeals to graph theorists, who are interested in sets of vertices of the n-cube that induce no 4-cycles. When restricting to just three middle levels, diamonds correspond to the only 4-cycles.

Note that these three middle levels each have size $\sim \binom{n}{\lfloor \frac{n}{2} \rfloor}$, so for a family $\mathcal{F}$ in these levels, $|\mathcal{F}|/\binom{n}{\lfloor \frac{n}{2} \rfloor}$ is essentially just the sum, over the three levels, of the proportions of elements in $\mathcal{F}$. Surely, we can manage this narrower problem on just three levels? In 2012 Axenovich, Manske, and Martin [2] reduced the upper bound for three-level diamond-free families to below the current best bound of $2.25\binom{n}{\lfloor \frac{n}{2} \rfloor}$ for general diamond-free families. They also observed that a bound on this problem also bounds the size of families in which the sets have just three sizes, not necessarily the middle three. Elegant arguments in 2013 by Manske and Shen [40] reduce it further. More recently Balogh, Hu, Lidický, and Liu (2014) employed flag algebras, a sophisticated method for associating inequalities to certain problems. The resulting semidefinite program then requires considerable computer time to produce good bounds. This method has proven successful in Turán theory [47, 3].

Theorem 3.2 *Let $\mathcal{F}$ be a $\mathcal{D}_2$-free family of subsets in $\mathcal{B}(n,3)$, that is, every set in $\mathcal{F}$ has one of the three middle sizes in n. Then we have*

(1)[2] $|\mathcal{F}| \leq 2.20\binom{n}{\lfloor \frac{n}{2} \rfloor}$,

(2)[40] $|\mathcal{F}| \leq 2.16\binom{n}{\lfloor \frac{n}{2} \rfloor}$,

(3)[4] $|\mathcal{F}| \leq 2.15121\binom{n}{\lfloor \frac{n}{2} \rfloor}$.

We propose this even more restricted problem that remains open:

Problem 3.3 *Show that if family $\mathcal{F}$ contains $\sim 70\%$ of the subsets from each of the three middle sizes, then $\mathcal{F}$ must contain a diamond $\mathcal{D}_2$.*

Notice that $|\mathcal{F}| \sim 2.1\binom{n}{\lfloor \frac{n}{2} \rfloor}$, which is less than what we need to use the bound of Balogh *et al.* That we cannot yet solve this toy problem suggests we do not yet fully understand how multiple levels of the Boolean lattice–even just three levels–interrelate.

Let us mention here additional work on diamond-free families. Czabarka, Dutle, Johnston, and Székely [11] used random families arising from posets based on abelian groups to construct more examples of large diamond-free families. These families can be viewed as additional support for the conjectured asymptotic behavior.

Dove [15] devised examples of diamond-free families $\mathcal{F}$ for all even $n \geq 6$ with $|\mathcal{F}| > \Sigma(n,2)$. It means that $\mathrm{La}(n,\mathcal{D}_2)$ is not likely to have a simple formula for all large n (unlike, for instance, $\bowtie$). These examples are all contained in $\mathcal{B}(n,3)$, the middle three layers. The asymptotic behavior of $\mathrm{La}(n,\mathcal{D}_2)$ must be more complicated than for examples like $\bowtie$, so it is unlikely anyone can determine these values exactly.

Sarkis, Shahriari, and students [48] investigated diamond-free families in the subspace lattice, and discovered the analogous asymptotic problem is less formidable.

In a later section we shall say more about $\mathrm{La}(n,P)$ for general diamonds $\mathcal{D}_k$. For now, let us mention one simple way k-diamonds arise. Griggs noticed that early results suggest a dependence of $\mathrm{La}(n,P)$ on the height $h(P)$. Indeed, from the result of DeBonis-Katona on forbidding $\mathcal{K}_{r,s}$ it follows that for every poset P of height at most two, $\mathrm{La}(n,P)/\binom{n}{\lfloor \frac{n}{2} \rfloor} \leq 2 + o(1)$ as $n \to \infty$. Griggs wondered whether for general h there is a bound c_h such that $\mathrm{La}(n,P)/\binom{n}{\lfloor \frac{n}{2} \rfloor} \leq c_h$ for all posets P of height h. This turned out to be completely false: Jiang and Lu [23] independently noticed that while the k-diamonds $\mathcal{D}_k$ all have height 3, there is no bound on how large $\mathrm{La}(n,\mathcal{D}_k)/\binom{n}{\lfloor \frac{n}{2} \rfloor}$ can get. This is easily seen from the fact that the family of the the middle r levels, $\mathcal{B}(n,r)$, contains no diamond $\mathcal{D}_k$ for $k \geq 2^{r-1}-1$. It means that for any r, for large enough k, we have $\mathrm{La}(n,\mathcal{D}_k)/\binom{n}{\lfloor \frac{n}{2} \rfloor} > r$ for all large n.

4 Conjectured Asymptotics for General Posets

For some posets P, such as chains or the butterfly (when $n \geq 3$), $\mathrm{La}(n,P)$ is simply a sum of largest binomial coefficients in n. For many more posets P, this fails, but the asymptotic behavior of $\mathrm{La}(n, P)$ is nice. Examples of this include $\mathcal{V}, \mathcal{N}$, and $\mathcal{K}_{r,s}$. Of course, we have seen that for the diamond $\mathcal{D}_2$, even the asymptotic behavior remains unsettled.

In their study of a variety of posets, Griggs and Lu [24] came to believe that for every finite poset P, as $n \to \infty$, $\mathrm{La}(n, P)$ should be asymptotic to $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ times some integer depending on P. This same belief likely guided the early work of Katona and his collaborators, even if it was not explicitly stated. Even now, some years later, the conjecture continues to hold for all posets for which the asymptotic behavior is known.

For posets P Griggs and Lu introduced the notation

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

The conjectured limit imitates the Turán density $\pi(H)$ of k-graphs H, which concerns the maximum number $\mathrm{ex}(n, H)$ of edges in a k-graph on n vertices not containing H, $\pi(H) := \lim_{n\to\infty} \frac{\mathrm{ex}(n,H)}{\binom{n}{k}}$. A difference is that the hypergraph limit is known to exist in general, while proving existence remains open for $\pi(P)$. Also, the hypergraph limit is not integer in general, while the poset limit is conjectured to be integer.

When Griggs announced this conjecture at a 2008 conference, two of the experts present, M. Saks and P. Winkler, both pointed out a pattern that explains the known limiting values $\pi(P)$, having to do with finding poset P in Boolean lattices $\mathcal{B}_n$. Griggs and Lu [23] subsequently formulated it by introducing notation: For a poset P, let $e(P)$ denote the maximum k such that the union $\mathcal{B}(n, k)$ of k middle levels in $\mathcal{B}_n$ does not contain P as a subposet, no matter how large n is.

For instance, the butterfly poset is not contained in $\mathcal{B}(n, 3)$ when $n = 2$, since $\mathcal{B}(2, 3)$ is simply the diamond $\mathcal{B}_2$. Starting at $n = 3$, $\mathcal{B}(n, 3)$ does contain (many) butterflies. On the other hand, $\bowtie$ is not contained in the union of any two consecutive levels in the Boolean lattice. It means that $e(\bowtie) = 2$. Since $\Sigma(n, 2) \sim 2\binom{n}{\lfloor \frac{n}{2} \rfloor}$, it gives $\pi(\bowtie) = 2 = e(\bowtie)$.

In general, the family $\mathcal{B}(n, e)$ contains no P for $e = e(P)$. So if $\pi(P)$ exists, it must be at least $e(P)$. The Griggs-Lu Conjecture, announced in 2008 [cf. [23]], is that $\pi(P)$ exists, and its value is $e(P)$:

Conjecture 4.1 [24, 23] *For any poset P, the limit $\pi(P) := \lim_{n\to\infty} \frac{\mathrm{La}(n,P)}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}$ exists. Moreover, its value is the integer $e(P)$.*

Not long after the conjecture was announced, in 2009, Bukh proved the existence of $\pi(P)$ for every poset which has a tree (acyclic) Hasse diagram:

Theorem 4.1 [7] *For all tree posets T, $\pi(T) = h(T) - 1 = e(T)$.*

This may be the strongest result yet in support of the Griggs-Lu Conjecture, considering that it includes so many of the posets described already, such as chains, forks, batons, height 2 trees, and $\mathcal{J}$. The proof is an impressive display of probabilistic reasoning.

Moreover, Bukh's result can be used to get a good estimate for other posets. For example, it gives another way to derive $\pi(\bowtie) = 2$: We already have the easy lower bound from $\mathrm{La}(n, \bowtie) \geq \mathcal{B}(n, 2) \sim 2\binom{n}{\lfloor \frac{n}{2} \rfloor}$. For the upper bound, we view the butterfly as a subposet of $\mathcal{X}$, where $\mathcal{X}$ is the poset on five elements with partial order relations $x_1, x_2 \leq x_3 \leq x_4, x_5$. We have that $\mathrm{La}(n, \bowtie) \leq \mathrm{La}(n, \mathcal{X})$. Since the poset $\mathcal{X}$ is a tree of height 3, by Theorem 4.1, $\mathrm{La}(n, \mathcal{X}) = (2 + o(1))\binom{n}{\lfloor \frac{n}{2} \rfloor}$.

5 Lubell Function

We have the general lower bound $e(P)$ on $\pi(P)$, when it exists, directly from the definitions. To verify the conjecture, the key is to get the upper bound. The Lubell function is a tool that was introduced to do just that, at least for some posets. This method is an extension of the heart of Lubell's elegant proof [39] of Sperner's Theorem (the case that $P = \mathcal{P}_2$).

Let $\mathscr{C} := \mathscr{C}_n$ denote the collection of all $n!$ *full (maximal) chains*

$$\emptyset \subset \{i_1\} \subset \{i_1, i_2\} \subset \cdots \subset [n]$$

in the Boolean lattice $\mathcal{B}_n$. Fix a family $\mathcal{F} \subseteq 2^{[n]}$. We say the *height* of $\mathcal{F}$ is

$$h(\mathcal{F}) := \max_{\mathcal{C} \in \mathscr{C}} |\mathcal{F} \cap \mathcal{C}|.$$

It is the same as the height if $\mathcal{F}$ is viewed as a poset under inclusion. Following [23] we define the *Lubell function* of $\mathcal{F}$ by

$$\bar{h}(\mathcal{F}) = \bar{h}_n(\mathcal{F}) := \operatorname*{ave}_{\mathcal{C} \in \mathscr{C}_n} |\mathcal{F} \cap \mathcal{C}|.$$

The bar notation suggests what it is, the average number of times a random full chain meets $\mathcal{F}$, as compared to the height, which is the maximum number of times over all full chains. The Lubell function value bounds the size of a family (extending Lubell's argument for an antichain $\mathcal{F}$):

Lemma 5.1 [23] *Let $\mathcal{F}$ be a collection of subsets of $[n]$. Then $\bar{h}(\mathcal{F}) = \sum_{F \in \mathcal{F}} 1/\binom{n}{|F|} \geq |\mathcal{F}|/\binom{n}{\lfloor \frac{n}{2} \rfloor}$.*

Proof. Let us count the total number of times all full chains meet $\mathcal{F}$ in two different ways. Then the average will be

$$\frac{1}{n!} \sum_{\mathcal{C} \in \mathscr{C}} |\mathcal{F} \cap C| = \frac{1}{n!} \sum_{F \in \mathcal{F}} |F|!(n - |F|)! = \sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}}.$$

The last inequality in the lemma follows from the fact that $\binom{n}{\lfloor \frac{n}{2} \rfloor} \geq \binom{n}{k}$ for all k. $\square$

The Lemma tells us that the ratio we are interested in, $|\mathcal{F}|/\binom{n}{\lfloor\frac{n}{2}\rfloor}$, is bounded above by the Lubell function of $\mathcal{F}$. So upper bounds on the Lubell function for all P-free families $\mathcal{F}$ yield upper bounds on $\mathrm{La}(n,P)/\binom{n}{\lfloor\frac{n}{2}\rfloor}$, the ratio that is conjectured to converge to $e(P)$.

We can view $\bar{h}(\mathcal{F})$ as a weighted sum, where each set F has weight $1/\binom{n}{|F|}$. Given P and n, let us define $\lambda_n(P)$ to be $\max \bar{h}_n(\mathcal{F})$ over all P-free families $\mathcal{F}$ in $\mathcal{B}_n$. For a largest P-free family $\mathcal{F}$, we have $\mathrm{La}(n,P)/\binom{n}{\lfloor\frac{n}{2}\rfloor} = |\mathcal{F}|/\binom{n}{\lfloor\frac{n}{2}\rfloor} \le \lambda_n(P)$, since $1/\binom{n}{\lfloor\frac{n}{2}\rfloor}$ is the minimum possible weight among all subsets of $[n]$. Therefore, if $\pi(P)$ exists, then it cannot exceed $\lim_{n\to\infty}\lambda_n(P)$. Let us denote this last limit by $\lambda(P)$, if it exists. We have that when the two limits exist,

$$e(P) \le \pi(P) \le \lambda(P).$$

How can one estimate $\lambda_n(P)$? Because every poset P is a subposet of a chain on $|P|$ elements, a full chain cannot intersect a P-free family $\mathcal{F}$ more than $|P|-1$ times. Hence, the average number of times a random full chain meets $\mathcal{F}$ is at most $|P|-1$. This gives us a very rough bound, $\lambda_n(P) \le |P|-1$. To refine it, we may suitably partition the set $\mathscr{C}_n$ of full chains into blocks and find for each block the average number of times a random full chain meets $\mathcal{F}$, and then take the maximum over all the blocks. This will typically overestimate $\bar{h}_n(\mathcal{F})$. Surprisingly, for many posets there is an appropriate choice of partition [21] that provides a practical way to evaluate $\lambda_n(P)$.

For general P, the existence of $\lambda(P)$ remains an open question. Moreover, unlike the lower bound, there are examples P for which $\pi(P) < \lambda(P)$, and this gap can be extremely large. Nevertheless, there are also some "good posets" P for which $\bar{h}_n(\mathcal{F}) \le e(P)$ for any P-free family $\mathcal{F}$ and any n. For such a poset, not only is the value of $\mathrm{La}(n,P)$ determined, but so are the families achieving it. Griggs and Li [22] say such posets are *uniformly L-bounded*, where L stands for Lubell.

Theorem 5.2 [23] *Let P be a poset that is uniformly L-bounded. Let $e = e(P)$. Then for all n, $\mathrm{La}(n,P) = \Sigma(n,e)$, and so $\pi(P) = e(P)$. Moreover, if $\mathcal{F}$ is a P-free family of subsets of $[n]$ of maximum size, $\mathcal{F}$ must be the family $\mathcal{B}(n,e)$.*

We see that uniformly L-bounded posets satisfy Conjecture 4.1. In [22], one can see a constructive method to obtain many uniformly L-bounded posets.

One poset of interest that definitely is not uniformly L-bounded is the notorious diamond $\mathcal{D}_2$. In fact, there are examples of diamond-free families $\mathcal{F}_n$ of subsets of $[n]$ for which the Lubell function approaches 2.25 as $n \to \infty$ [23]. Consequently, the sequence $\lambda_n(\mathcal{D}_2)$, which is known to be nonincreasing in n by a simple averaging argument, has a limit $\lambda(\mathcal{D}_2) \ge 2.25$. It means that working purely with the Lubell function, as in [23], one cannot reduce the upper bound on $\pi(\mathcal{D}_2)$ any lower than 2.25.

Kramer *et al.* [35] used the Lubell function in their proof that brought the upper bound on $\pi(\mathcal{D}_2)$ down to the very same value, 2.25. However, they were able to restrict their consideration to diamond-free families of a particular type. As a consequence, the value of $\lambda(\mathcal{D}_2)$ remains open, with lower bound 2.25 and upper bound about 2.273 by the proof of [23].

So why should we still expect that $\pi(\mathcal{D}_2) = 2$? The families $\mathcal{F}_n$ we mentioned, with relatively large Lubell function values, are actually not very large. They consist of very small subsets (size at most 3), or else dually very large subsets (size at least $n-3$): Each set in $\mathcal{F}_n$ makes a rather large contribution to $\bar{h}(\mathcal{F}_n)$.

What happens with the k-diamonds $\mathcal{D}_k$ for general k? Do they get even harder to deal with? Surprisingly, Li discovered that for "almost all" values of k, including k as small as 3 or 4, $\mathcal{D}_k$ is uniformly L-bounded, so satisfies the Griggs-Lu Conjecture, and we know the exact values of $\mathrm{La}(n, P)$ and the extremal families. For other values of k it remains a great challenge to determine $\mathrm{La}(n, P)$:

Theorem 5.3 [23, 36] *Let $k \geq 2$, and define $m := \lceil \log_2(k+2) \rceil$.*

(1) If $k \in [2^{m-1} - 1, 2^m - \binom{m}{\lfloor \frac{m}{2} \rfloor} - 1]$, then $\mathcal{D}_k$ is uniformly L-bounded, and $\pi(\mathcal{D}_k) = e(\mathcal{D}_k) = m$.

(2) If $k \in [2^m - \binom{m}{\lfloor \frac{m}{2} \rfloor}, 2^m - 2]$, then, if $\pi(\mathcal{D}_k)$ exists,

$$m = e(\mathcal{D}_k) \leq \pi(\mathcal{D}_k) \leq m + 1 - \frac{2^m - k - 1}{\binom{m}{\lfloor \frac{m}{2} \rfloor}} < m + 1.$$

Another class of posets that is uniformly L-bounded is the set of suspensions of paths of distinct lengths. Precisely, for $k \geq 1$ let $l_1 \geq \cdots \geq l_k \geq 3$, and define the *harp poset* $\mathcal{H}(l_1, \ldots, l_k)$ to consist of paths $\mathcal{P}_{l_1}, \ldots, \mathcal{P}_{l_k}$ with their top elements identified and their bottom elements identified. For instance, in this notation we have $\mathcal{D}_k$ is the harp $\mathcal{H}(3, \ldots, 3)$ where there are k 3's. Then Griggs, Li, and Lu proved:

Theorem 5.4 [23] *If $l_1 > \cdots > l_k \geq 3$, then $\mathrm{La}(n, \mathcal{H}(l_1, \ldots, l_k))$ is uniformly L-bounded, and $\pi = e = l_1 - 1$.*

For example, the unranked poset on five elements called N_5 by lattice theorists is the harp $\mathcal{H}(4, 3)$. By the theorem above, it is uniformly L-bounded, so we know that $\mathrm{La}(n, N_5) = \Sigma(n, 3)$ for all n.

The vast majority of subsets of $[n]$ are close to the middle value in size. There are comparatively few subsets that are small or large. More precisely, Shannon's Theorem [1, p.256] gives

$$\sum_{i=0}^{[\alpha n]} \binom{n}{i} = o\left(\binom{n}{\lfloor \frac{n}{2} \rfloor}\right),$$

for any positive constant $\alpha < 1/2$, showing there are few small sets. By taking complements there are similarly few large sets.

On the other hand, sets with either few or many elements make major contributions to the weighted sum that is the Lubell function. In using the Lubell function to bound $|\mathcal{F}|/\binom{n}{\lfloor \frac{n}{2} \rfloor}$ asymptotically for P-free families, it is enough to consider such families $\mathcal{F}$ that contain no small or large sets, thereby avoiding the overly heavy contributions of small and large sets to $\bar{h}(\mathcal{F})$. A full hierarchy of poset properties has been devised by Griggs and Li [22], involving the Lubell function, to take advantage of this observation. Each of the properties imply that $\pi(P) = e(P)$, though the values of $\mathrm{La}(n, P)$ may not be determined precisely as with uniformly L-bounded posets.

We say for integer m that poset P is *m-L-bounded*, if for every n every P-free family $\mathcal{F} \subseteq \mathcal{B}_n$, containing only sets of sizes between m and $n-m$, satisfies $\bar{h}_n(\mathcal{F}) \le e(P)$. For $m = 0$, these are uniformly L-bounded posets. The butterfly poset is not uniformly L-bounded, but it is 1-L-bounded. The m-L-bounded posets satisfy Conjecture 4.1.

As an application Griggs and Li [22] introduced the class of fans, where a *fan poset* is the join of paths, that is, paths (chains) that are identified at the bottom. Fans all satisfy some Lubell boundedness condition, so that $\pi(P) = e(P)$, which we already knew by Bukh's Tree Theorem. However, now we know more. For instance, if there is just one path, or if the path lengths are distinct with the longest path at least 2 longer than the others, the fan is uniformly L-bounded. In this case we know $\mathrm{La}(n, P)$ exactly, as well as the extremal families.

Griggs and Li [22] presented a series of results describing how to construct many more posets satisfying Lubell boundedness properties (and thus the $\pi(P) = e(P)$ conjecture). It is surprising one can build such complicated posets that satisfy the conjecture, when we still seem so far from verifying it for the diamond $\mathcal{D}_2$.

6 Forbidden Induced Subposets

Up to this point, we only considered families that forbid a subposet P in the weak sense, meaning that we are also excluding all posets that contain P. It is perhaps equally natural to investigate what happens if we exclude only families of subsets ordered exactly as P is. A poset P_1 contains poset P_2 as an *induced subposet*, if there is an injection f (a strong embedding) from P_2 to P_1, such that for $x, y \in P_2$, $x \le_{P_2} y$ if and only if $f(x) \le_{P_1} f(y)$. For instance, the diamond $\mathcal{D}_2$ is a subposet of the chain $\mathcal{P}_4$, but it is not an induced subposet of any chain, no matter how long.

The notation, $\mathrm{La}^{\#}(n, P)$, introduced by Carroll and Katona [9], is the maximum size of a family $\mathcal{F}$ of subsets of $[n]$ not containing the poset P as an induced subposet. Clearly, $\mathrm{La}(n, P) \le \mathrm{La}^{\#}(n, P)$, since it is more restrictive to forbid P in the weak sense.

Let us begin again with chains. Since a family $\mathcal{F}$ contains the chain $\mathcal{P}_k$ in the usual weak sense if and only if it contains it in the strong sense, we have $\mathrm{La}(n, \mathcal{P}_k) = \mathrm{La}^{\#}(n, \mathcal{P}_k)$, which is $\Sigma(n, k-1)$, by Erdős's Theorem on chains.

Next, consider $\mathcal{V} = \mathcal{V}_2$, which is what Carroll and Katona did (2008):

Theorem 6.1 [9]

$$1 + \frac{1}{n} + \Omega\left(\frac{1}{n^2}\right) \le \frac{\mathrm{La}^{\#}(n, \mathcal{V}_2)}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \le 1 + \frac{2}{n} + O\left(\frac{1}{n^2}\right).$$

So far, the results are similar to usual weak containment. However, in general it should be much easier to build families that only avoid containing P as an *induced* subposet, since in order for $\mathcal{F}$ to contain an induced copy of P, not only must we have the containment relations intact, we must also have the non-containments.

Indeed, observe that since a poset P is always a subposet of the chain $\mathcal{P}_{|P|}$, it means that

$$\mathrm{La}(n,P) \le \mathrm{La}(n,\mathcal{P}_{|P|}) = \Sigma(n,|P|-1) \sim (|P|-1)\binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

By contrast, for general P it is not at all clear how to bound $\mathrm{La}^{\#}(n,P)/\binom{n}{\lfloor \frac{n}{2} \rfloor}$ for all n. Is it even bounded in general? Lu and Milans [38] have conjectured that it is bounded for every poset P (and it is reported Katona also made this conjecture).

In more detail, for the induced subposet problem, Lu and Milans defined the *Turán threshold*

$$\pi^*(P) := \limsup_{n\to\infty} \frac{\mathrm{La}^{\#}(n,P)}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}$$

and the *Lubell threshold*

$$\lambda^*(P) := \limsup_{n\to\infty} \{\bar{h}_n(\mathcal{F}) \mid P \text{ is not as an induced subposet of } \mathcal{F}\}.$$

They conjectured that the two thresholds are finite for all posets. Clearly, if $\lambda^*(P)$ is finite, so is $\pi^*(P)$, since $\pi^*(P) \le \lambda^*(P)$. Lu and Milans verified that $\lambda^*(P)$ is finite for series-parallel posets and for posets of height at most 2. These innocent-looking claims stymied the experts for some time. (We return to this shortly.)

Given the challenge of working with $\mathrm{La}^{\#}(n,P)$, it is then remarkable that Boehnlein and Jiang (2012) managed to extend Bukh's methods to determine the asymptotics of all tree posets:

Theorem 6.2 [5] *Let T be a tree poset of height $h \ge 2$. Then*

$$\mathrm{La}^{\#}(n,T) = (h-1)\binom{n}{\lfloor \frac{n}{2} \rfloor}(1+o(1)).$$

We see that $\pi^*(T) = h-1$ for tree posets of height h.

Next compare $\mathrm{La}^{\#}(n,P)$ with $\mathrm{La}(n,P)$, which in general is smaller. For tree posets P, we have that $\lim_{n\to\infty} \mathrm{La}^{\#}(n,P)/\mathrm{La}(n,P) = 1$. On the other hand, there are examples of posets P such that $\lim_{n\to\infty} \mathrm{La}^{\#}(n,P)/\mathrm{La}(n,P)$ is arbitrarily large [5]. For example, let H_m be a poset consisting of elements $x_1,\ldots,x_m$ and $y_1,\ldots,y_m$ with orderings $x_i \le y_j$ for $i \le j \le m$. This is a poset of height 2, which is contained in $\mathcal{K}_{m,m}$ as a non-induced poset. On the other hand, when $m \ge 3$, H_m contains the butterfly as a subposet. Thus, by Theorems 2.3 and 2.5, $\mathrm{La}(n,H_m) \sim 2\binom{n}{\lfloor \frac{n}{2} \rfloor}$. However, one can show that $\mathrm{La}^{\#}(n,H_m) \ge (m-1-o(1))\binom{n}{\lfloor \frac{n}{2} \rfloor}$, which is larger by a factor of $\frac{m-1}{2}$ than $\mathrm{La}(n,H_m)$.

For complete two-level posets, Patkós (2015) gave the following estimates of $\mathrm{La}^{\#}(n,P)$:

Theorem 6.3 [46] *For integers $r, s \geq 2$, we have*

$$\Sigma(n,2) + \left(\frac{r+s-2}{n} - O_{r,s}\left(\frac{1}{n^2}\right)\right)\binom{n}{\lfloor \frac{n}{2} \rfloor}$$
$$\leq \quad \mathrm{La}^{\#}(n, K_{r,s}) \leq \left(2 + \frac{2(r+s-2)}{n} + o\left(\frac{1}{n}\right)\right)\binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

It is interesting that while H_m is a subposet of $\mathcal{K}_{m,m}$, we have $\mathrm{La}^{\#}(n, H_m) \geq \mathrm{La}^{\#}(n, \mathcal{K}_{m,m})$.

What about the notorious diamond poset $\mathcal{D}_2$? Lu and Milans (2014) gave the best result to date:

Theorem 6.4 [38] *As $n \to \infty$, $\mathrm{La}^{\#}(n, \mathcal{D}_2) \leq (2.583 + o(1))\binom{n}{\lfloor \frac{n}{2} \rfloor}$.*

Later in 2014 there was a breakthrough on the induced subposet problem, when the boundedness conjecture of Katona and of Lu and Milans was proven by Methuku and Pálvőlgyi. They brought the ideas and results from the area of forbidden hypermatrices [33, 41] to the problem and established the next theorem. It is interesting that the parameter, the order dimension of P, has never appeared in the theory before.

Theorem 6.5 [43] *For every poset P, there is a $C_P = 2^d K$ such that $\mathrm{La}^{\#}(n, P) \leq C_P\binom{n}{\lfloor \frac{n}{2} \rfloor}$, where d is the order dimension of P and K is a constant obtained from forbidden hypermatrix problems.*

It means that $\pi^*(P) \leq C_P$. No sooner did Methuku and Pálvőlgyi prove Theorem 6.5, than the (stronger) Lu-Milans conjecture that $\lambda^*(P)$ is finite for all P was proven by Méroueh with the help of the probablistic tools.

Theorem 6.6 [42] *For every poset P, there exists $c(P)$ such that if $\mathcal{F} \subset 2^{[n]}$ for some $n \in \mathbb{N}$ and if $\mathcal{F}$ does not contain P as an induced subposet, then $\bar{h}(\mathcal{F}) \leq c(P)$.*

7 General Upper Bounds

Erdős's result, Theorem 1.2, provides an upper bound of $\Sigma(n, k-1)$ on $\mathrm{La}(n, P)$ for any poset P on k elements. This bound is far from accurate when the *width*, the largest size of an antichain in P, is large. In this section, we present ideas for finding upper bounds on $\mathrm{La}(n, P)$ using the invariants of P.

Recall that the Lubell function for a family $\mathcal{F}$ is the average number of times that a full chain $\mathcal{C}$ meets $\mathcal{F}$. We can also count the average number of times other families meet $\mathcal{F}$. Bursci and Nagy (2013) called a family of sets in $\mathcal{B}_n$ a *double chain* $\mathcal{D}$, as follows. It consists of a full chain

$$\mathcal{C} = \{I_0 \subset I_1 \subset \cdots \subset I_n\},$$

where $|I_i| = i$ for all i, called the primary chain, together with sets $S_i = I_{i-1} \cup I_{i+1} \backslash I_i$, $1 \leq i \leq n-1$. The number of double chains containing a specific set is twice the number of full chains containing it, for every subset besides $\emptyset, [n]$. Using a double counting argument as in Lubell's method, there is a weighted sum of sets in a P-free family:

$$\frac{1}{n!}\sum_{\substack{F\in\mathcal{F}\\F\neq\emptyset,[n]}} 2|F|!(n-|F|)! \leq \frac{1}{n!}\sum_{\mathcal{D}}|\mathcal{D}\cap\mathcal{F}| \leq (|P|+h(P)-2). \tag{1}$$

This leads to an upper bound that involves not only the size, but also the height of P:

Theorem 7.1 [8] *For any poset P,*

$$\mathrm{La}(n,P) \leq \left(\frac{|P|+h(P)-2}{2}\right)\binom{n}{\lfloor\frac{n}{2}\rfloor}.$$

Note that the height of a poset is at most the size. So, this result indeed improves Theorem 1.2 for general P. While Erdős's upper bound is only tight for chains, there is a series of posets P with $\mathrm{La}(n,P)$ reaching the upper bound, one of them being the butterfly.

Chen and Li (2014) introduced a new type of set family, called *k-linkage*, with a parameter k. They derived a more general inequality that reduces to the previous one when $k=1$:

Theorem 7.2 [10] *Given a poset P and $k\geq 1$, for sufficiently large n,*

$$\mathrm{La}(n,P) \leq \frac{1}{k+1}\left(|P|+\frac{1}{2}(k^2+3k-2)(h(P)-1)-1\right)\binom{n}{\lfloor\frac{n}{2}\rfloor}.$$

By setting $k=\lfloor\sqrt{|P|/h(P)}\rfloor$, it implies $\mathrm{La}(n,P)=O(\sqrt{|P|h(P)})\binom{n}{\lfloor\frac{n}{2}\rfloor}$.

The next inequality is proved by Grosz, Methuku, and Tompkins [26] using the concept of a *k-interval chain* $\mathcal{C}_k$, which is a union of a full chain $\emptyset\subset\{i_1\}\subset\{i_1,i_2\}\subset\cdots\subset[n]$ and the collection of sets S satisfying $\{i_1,\ldots,i_m\}\subset S\subset\{i_1,\ldots,i_{m+k}\}$ for some m. When $k=2$, a k-interval chain is a double chain $\mathcal{D}$. Counting $|\mathcal{C}_k\cap\mathcal{F}|$ for a P-free family $\mathcal{F}$ is technical and complicated, but it gives a better estimate for $\mathrm{La}(n,P)$.

Theorem 7.3 [26] *Given a poset P, for any integer $k\geq 2$, it holds that*

$$\mathrm{La}(n,P) \leq \frac{1}{2^{k-1}}\left(|P|+(3k-5)2^{k-2}h(P)-1)-1\right)\binom{n}{\lfloor\frac{n}{2}\rfloor}.$$

As in Theorem 7.2, we get

$$\mathrm{La}(n,P)=O\left(h(P)\log_2\left(\frac{|P|}{h(P)}+2\right)\right)\binom{n}{\lfloor\frac{n}{2}\rfloor}$$

by an appropriate choice of k. On the other hand, Grosz *et al.* also showed that for some posets P,

$$\mathrm{La}(n,P) \geq \left((h(P)-2)\log_2\left(\frac{|P|}{h(P)}\right)\right)\binom{n}{\lfloor\frac{n}{2}\rfloor}.$$

8 Packing posets

In their pioneering paper on forbidden subposets from 1983, Katona and Tarján [32] proposed the more general problem of forbidding a family. In our notation, given a collection of posets $\{Q_j\}_{j\geq 1}$ (possibly infinite), we seek the maximum size $\mathrm{La}(n, \{Q_j\}_{j\geq 1})$ of a family $\mathcal{F}$ of subsets of $[n]$ that contains no copy of any poset Q_j in the collection. Katona and Tarján, and independently but much later, Dove and Griggs [16] solved an interesting instance of this problem. Recall that Λ is the three-element poset that is the dual of $\mathcal{V} = \mathcal{V}_2$.

Theorem 8.1 [32, 16] *For all n,*

$$\mathrm{La}(n, \{\mathcal{V}, \Lambda\}) = 2\binom{n-1}{\lfloor \frac{n-1}{2} \rfloor}.$$

Dove and Griggs obtained this result by converting it to a packing problem. Instead of trying to pick many subsets of $[n]$ that avoid containing a poset P, we take the entire Boolean lattice $\mathcal{B}_n$ and seek to pack in as many unrelated copies of P as possible. Here, a *copy of* P is meant in the weak sense as usual. So if $\mathcal{F}$ avoids any $\mathcal{V}$ or Λ, it means that $\mathcal{F}$ can be decomposed into unrelated chains, each of size 2 or 1, and this is a problem that can be completely solved.

It depends mainly on how many unrelated copies of the chain $\mathcal{P}_2$ can be packed into $\mathcal{B}_n$. More precisely, following Dove and Griggs [16], let the packing number $\mathrm{Pa}(n, P)$ denote the maximum size of a family of subsets $\mathcal{F}$ constructed from pairwise unrelated copies of P in $\mathcal{B}_n$. Here $\mathcal{F}$ is the union of (weak) copies of P such that no two elements of different copies are equal or even related (by inclusion). We measure the size of such a family, rather than the number of copies of P. The measures are equivalent, of course, but later we will describe a more general model, where we build $\mathcal{F}$ out of unrelated copies of posets P_i coming from a collection. In this case, the family size is more important than the number of copies.

Here is the determination of $\mathrm{Pa}(n, P)$ when P is a chain $\mathcal{P}_k$. It was obtained in 1984 by Griggs, Stahl, and Trotter [25]. It can also be deduced from an earlier paper (1965) of Bollobás [6].

Theorem 8.2 [25] *The packing number for k-paths is $\mathrm{Pa}(n, \mathcal{P}_k) = k\binom{n-(k-1)}{\lfloor \frac{n-(k-1)}{2} \rfloor}$. For fixed k, as n goes to infinity, this is asymptotic to $\frac{k}{2^{k-1}}\binom{n}{\lfloor \frac{n}{2} \rfloor}$.*

It is interesting that the packing number for k-paths (chains) is asymptotic to a constant times $\binom{n}{\lfloor \frac{n}{2} \rfloor}$, the same behavior as for $\mathrm{La}(n, \mathcal{P}_k)$, though with a different constant. For general posets P, the asymptotics of $\mathrm{La}(n, P)$ remain rather wide open. It is then surprising that it is possible to determine the leading asymptotic behavior of the packing number $\mathrm{Pa}(n, P)$ for general posets P. Dove and Griggs asked and solved this question in 2013. Only later did they learn that the same problem had been posed first by Katona (2010) [29]. Moreover, Katona and Nagy also succeeded in solving it, at about the same time in 2013. The papers by the two teams were published together (2015) [16, 31].

Both teams formulated their results using convexity. A family $\mathcal{F} \subseteq \mathcal{B}_n$ generates an *ideal* (or *down-set*) and a *filter* (or *up-set*) denoted as follows:

$$D(\mathcal{F}) = \{S \in \mathcal{B}_n | S \subseteq A \text{ for some } A \in \mathcal{F}\}, \text{ and}$$

$$U(\mathcal{F}) = \{S \in \mathcal{B}_n | A \subseteq S \text{ for some } A \in \mathcal{F}\}.$$

Define the *convex closure operator* on $\mathcal{F}$ to be $\overline{\mathcal{F}} := D(\mathcal{F}) \cap U(\mathcal{F})$. Note that $\mathcal{F} \subseteq \overline{\mathcal{F}}$. It is important that if two subset families are unrelated, then their closures must be unrelated as well. If a family $\mathcal{F}$ satisfies $\mathcal{F} = \overline{\mathcal{F}}$, we say $\mathcal{F}$ is *convex*.

The asymptotic determination of $\mathrm{Pa}(n, P)$ depends on how P can be weakly embedded into a Boolean lattice. There exists a minimum size of the closure of $f(P)$ over all weak embeddings f of P into $\mathcal{B}_k$ over all k. Denote this minimum by $c(P)$.

Theorem 8.3 [16, 31] *For any poset P, as $n \to \infty$,* $\mathrm{Pa}(n, P) \sim \dfrac{|P|}{c(P)} \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

This result can be proven using Lubell-style chain averaging methods, extending the arguments used to prove Theorem 8.2.

Dove and Griggs expanded their study to allow packing posets from a collection. Specifically, let $\mathrm{Pa}(n, \{P_i\}_{i\geq 1})$ denote the maximum size of a family in $\mathcal{B}_n$ constructed from pairwise unrelated (weak) copies of posets, each chosen from a given collection of posets $\{P_i\}_{i\geq 1}$ (possibly infinite). Note that the sizes of the posets P_i can vary. Dove and Griggs showed that as n goes to infinity,

$$\mathrm{Pa}(n, \{P_1, P_2, \ldots, P_k\}) \sim \max_{1\leq i\leq k} \left(\frac{|P_i|}{c(P_i)} \right) \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

Similarly, one can investigate families built from *induced* copies of given posets. Denote the maximum size of a family in $\mathcal{B}_n$ constructed from induced copies of P as $\mathrm{Pa}^*(n, P)$. We can also define the more general quantity $\mathrm{Pa}^*(n, \{P_i\}_{i\geq 1})$. Now, for any collection $\{Q_j\}_{j\geq 1}$, the forbidden poset maximum, $\mathrm{La}(n, \{Q_j\}_{j\geq 1})$, is equivalent to $\mathrm{Pa}^*(n, \{P_i\}_{i\geq 1})$, where $\{P_i\}_{i\geq 1}$ is the collection of all possible connected posets that do not contain any of the posets in $\{Q_j\}_{j\geq 1}$ as a subposet. For instance, $\mathrm{La}(n, \mathcal{V}) = \mathrm{Pa}^*(n, \{\Lambda_i\}_{i\geq 0})$. So the problem of determining $\mathrm{Pa}^*(n, \{P_i\}_{i\geq 1})$ can be viewed as more general than the $\mathrm{La}(n, \{Q_j\}_{j\geq 1})$ problem this survey is addressing.

Both teams independently formulated and solved the analogue of the result above for induced subposets. We need to define $c^*(P)$ as the minimum size of the closure of a strong embedding of P in $\mathcal{B}_n$ over all possible n. In general, $c^*(P) \neq c(P)$. Here is the main theorem for packing induced copies of P.

Theorem 8.4 [16, 31] *For any poset P, as $n \to \infty$,* $\mathrm{Pa}^*(n, P) \sim \dfrac{|P|}{c^*(P)} \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

9 Supersaturation

There has been some effort to obtain results in this area of "supersaturation" or "Erdős-Rademacher type" which concern how many copies of poset P there must be in every family $\mathcal{F} \subseteq \mathcal{B}_n$ of given size greater than $\mathrm{La}(n, P)$.

Kleitman [34] solved this problem in 1966 for the case $P = \mathcal{P}_2$, verifying a conjecture of Erdős-Katona that the minimum number of copies is attained by picking $\mathcal{F}$ to consist of subsets as close to the middle value as possible. He conjectured that for any k, the same families of subsets closest to the middle would achieve the minimum number of copies of chains $\mathcal{P}_k$ for any k.

This conjecture remains open for general k. However, for families of size not much larger than $\mathrm{La}(n, \mathcal{P}_k) = \Sigma(n, k-1)$, it has been verified independently by Dove, Griggs, Kang, and Séreni and by Das, Gan, and Sudakov. Using different approaches, both groups proved:

Theorem 9.1 [12, 17] *If a family $\mathcal{F}$ of subsets of $[n]$ satisfies $|\mathcal{F}| = \Sigma(n, k-1) + x$, then there must be at least*

$$x \cdot \prod_{i=1}^{k-1} \left(\left\lfloor \frac{n+k}{2} \right\rfloor - i + 1 \right)$$

copies of $\mathcal{P}_k$ in $\mathcal{F}$.

Note that this product

$$\prod_{i=1}^{k-1} \left(\left\lfloor \frac{n+k}{2} \right\rfloor - i + 1 \right)$$

is the number of copies of $\mathcal{P}_k$ contained in $\mathcal{B}(n, k)$, with one endpoint of the chain being a particular set in the kth middle rank (of ranks in $[0, n]$). Thus, the family that consists of $k-1$ middle ranks $\mathcal{B}(n, k-1)$ and x sets from the kth middle rank witnesses that the bound above is tight for

$$x \leq \binom{n}{\left\lfloor \frac{n}{2} \right\rfloor + (-1)^k \left\lfloor \frac{k}{2} \right\rfloor}.$$

Dove's dissertation [15] further addresses this problem, extending the theorem above to the lattice of subsets of a multiset (the divisor lattice).

The first result we know that goes beyond chains considers P to be the butterfly poset. Recall $\mathrm{La}(n, P) = \Sigma(n, 2)$. In 2014 Patkós (2014) [45] considered families of $\Sigma(n, 2) + x$ subsets of $[n]$, $x > 0$. For small values of x, similar to the result above for chains P, he determined the minimum possible number of butterflies. Beyond that, he found a general bound that is asymptotically best-possible.

10 Future Study

There are posets P for which we can readily determine $\mathrm{La}(n, P)$ for all n. In most cases, $\mathrm{La}(n, P)$ is simply a sum of binomial coefficients, often of the form $\Sigma(n, k)$ for some k.

It may even be possible to describe all of the extremal P-free families. Examples include chains $\mathcal{P}_k$, the butterfly $\bowtie$, and diamonds $\mathcal{D}_3, \mathcal{D}_4$.

Then there are posets P such as $\mathcal{V}$, for which it seems to be very hard to determine $\mathrm{La}(n, P)$, but for which the asymptotic behavior $\pi(P)$ is known.

However, there seem to be many posets P for which even the asymptotic behavior is challenging or beyond current methods. The most notable example is the diamond $\mathcal{D}_2$. Other examples include the 5-diamond $\mathcal{D}_5$, and the crowns $\mathcal{O}_6, \mathcal{O}_{10}$.

Beyond trying to find $\mathrm{La}(n, P)$, or at least $\pi(P)$, how can we tell which posets we can solve and which are going to difficult (or hopeless)?

Many researchers continue to be interested in the diamond-free family problem, particularly for families restricted to the middle three levels. We have learned that the upper bound on $\pi(\mathcal{D}_2)$ provided by the Lubell function can only go down to 2.25, which is still well above the conjectured upper bound of 2. If the sets in the family $\mathcal{F}$ are excluded from the smallest and largest sizes–which won't really affect the asymptotics of $|\mathcal{F}|/\binom{n}{\lfloor \frac{n}{2} \rfloor}$ since asymptotically few subsets are very small or very large–it is suspected the Lubell bound can be brought down to the conjectured 2. However, no one has been able to reduce the Lubell bound with such size restrictions.

A smaller problem that remains open is to prove that the Lubell limit $\lambda(\mathcal{D}_2) = 2.25$. We have the lower bound, and know the limit exists, but the upper bound remains around 2.27.

Flag algebra arguments (and related computations) have led to improved bounds for the diamond poset $\mathcal{D}_2$. Can they give the conjectured value of 2, or solve other open subposet problems, as they have in Turán theory?

After the diamond $\mathcal{D}_2$, the most intriguing posets for which the existence of $\pi(P)$ remains open may be the crowns $\mathcal{O}_6$ and $\mathcal{O}_{10}$.

Probability theory has figured into Lubell function arguments and in the proof of Bukh's Tree Theorem. It seems likely that we can do more with probability to answer the asymptotic questions raised in this field.

It feels like we have only gotten started in the study of induced subposet problems. The recent proof of the conjecture of Katona and Lu-Milans, that the Turán threshold $\pi^*(P)$ is finite for all posets P, is encouraging. It remains open to prove the conjecture of Lu-Milans that the Lubell threshold $\lambda^*(P)$ is finite for all posets P.

This survey has not done justice to all of the wonderful ideas developed in the area, but at least we have collected all of the references to date we are aware of, and the interested reader is encouraged to study them. This continues to be a lively research area, and there is a long way to go!

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, 3rd edition Wiley, New York (2008).

[2] M. Axenovich, J. Manske, and R. Martin, Q_2-free families in the Boolean lattice, *Order* **29**, (2012), 177–191.

[3] R. Baber and J. Talbot, Hypergraphs do jump, *Combinatorics, Probability, and Computing* **20** (2011), 161–171.

[4] J. Balogh, P. Hu, B. Lidický, and H. Liu, Upper bounds on the size of 4- and 6-cycle-free subgraphs of the hypercube, *European Journal of Combinatorics* **35** 2014, 75-85.

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

[6] B. Bollobás, On generalized graphs, *Acta Math. Acad. Sci. Hung.* **16** (1965), 447-452.

[7] B. Bukh, Set families with a forbidden poset, *Elect. J. Combin.* **16** (2009), R142, 11p.

[8] P. Burcsi and D. T. Nagy, The method of double chains for largest families with excluded subposets, *Elect. J. Graph Theory and Applications*, Vol **1** (2013), 40–49.

[9] T. Carroll and G. O. H. Katona, Bounds on maximal families of sets not containing three sets with $A \cup B \subset C, A \not\subset B$, *Order* **25** (2008) 229–236.

[10] H.-B. Chen and W.-T. Li, A note on the largest size of families of sets with a forbidden poset, *Order* **31** (2014), 137–142.

[11] É. Czabarka, A. Dutle, T. Johnston, and L. A. Székely, Abelian groups yield many large families for the diamond problem, *Eur. J. Math.* **1** (2015), 320–328.

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

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

[14] A. DeBonis, G. O. H. Katona, and K. J. Swanepoel, Largest family without $A \cup B \subset C \cap D$, *J. Combinatorial Theory (Ser. A)* **111** (2005), 331–336.

[15] A. P. Dove, Generalizations of Sperner's Theorem: Packing Posets, Families Forbidding Posets, and Supersaturation, Ph.D. dissertation, University of South Carolina, 2013.

[16] A. P. Dove and J. R. Griggs, Packing posets in the Boolean lattice, *Order* **32** (2015), 429–438.

[17] A. P. Dove, J. R. Griggs, R. J. Kang, and J.-S. Sereni, Supersaturation in the Boolean Lattice, *INTEGERS*, **14A** (2014).

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

[19] R. L. Graham and N. J. A. Sloane, Lower bounds for constant weight codes, *IEEE IT* bf 26 (1980), 37-43.

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

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

[22] J. R. Griggs and W.-T. Li, Poset-free families and Lubell-boundedness, *J. Combinatorial Theory (Ser. A)* **134** (2015), 166–187.

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

[24] J. R. Griggs and L. Lu, On families of subsets with a forbidden subposet, *Combin. Probab. Comput.* **18** (2009), 731–748.

[25] J. R. Griggs, J. Stahl, and W. T. Trotter, Jr., A Sperner Theorem on Unrelated Chains of Subsets, *J. Combinatorial Theory (Ser. A)* **36** (1984), 124–127.

[26] D. Grósz, A. Methuku, and C. Tompkins, An improvement of the general bound on the largest family of subsets avoiding a subposet P, *Order* **34** (2017), 113–125.

[27] D. Grósz, A. Methuku, and C. Tompkins, An upper bound on the size of diamond-free families of sets avoiding a subposet P, ArXiv:1601.06332

[28] G. O. H. Katona, Forbidden inclusion patterns in the families of subsets (introducing a method), in *Horizons of Combinatorics*, Bolyai Society Mathematical Studies, **17**, Bolyai Mathematical Society, Budapest and Springer-Verlag, 2008, pp. 119–140.

[29] G. O. H. Katona, On the maximum number of incomparable copies of a given poset in a family of subsets, Lecture, International Conference on Recent Trends in Graph Theory and Combinatorics, Cochin University of Science and Technology, India (2010).

[30] G. O. H. Katona, Sperner type theorems with excluded subposets, *Discrete Applied Mathematics* **161** (2013), 1251–1258.

[31] G. O. H. Katona and D. Nagy, Incomparable Copies of a Poset in the Boolean Lattice *Order* **32** (2015), 419–427.

[32] G. O. H. Katona and T. G. Tarján, Extremal problems with excluded subgraphs in the n-cube, in: M. Borowiecki, J. W. Kennedy, and M. M. Sysło (eds.) **Graph Theory**, Łagów, 1981, *Lecture Notes in Math.*, **1018** 84–93, Springer, Berlin Heidelberg New York Tokyo, 1983.

[33] M. Klazer and A. Marcus, Extensions of the linear bound in the F uredi-Hajnal conjecture, *Adv. in Appl. Math.* **38** (2006), 258–266.

[34] D. Kleitman, A conjecture of Erdős-Katona on commensurable pairs among subsets of an n-set, *Theory of Graphs, Proceedings of the Colloquium Held at Tihany, Hungary, September 1966 (P. Erdős and G. O. H. Katona, eds.), Academic Press, New York*, (1968) 215-V218. Theory of Graphs, Proc. Colloq., Tihany (1966), 215 V 218.

[35] L. Kramer, R. R. Martin, and M. Young, On diamond-free subposets of the Boolean lattice, *J. Combinatorial Theory (Ser. A)* **120** (2013), 545–560.

[36] W.-T. Li, Extremal problems on families of subsets with forbidden subposets, Ph.D. dissertation, University of South Carolina, 2011.

[37] L. Lu, On crown-free famililes of subsets, *J. Combinatorial Theory (Ser. A)* **126** (2014), 216–231.

[38] L. Lu and K. Milans, Set families with forbidden subposets, *J. Combinatorial Theory (Ser. A)* **136** (2015), 126–142.

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

[40] J. Manske and J. Shen, Three layer Q_2-free families in the Boolean lattice, *Order* **30** (2013), 585–592.

[41] A. Marcus and G. Tardos, Exclided permutation matrices and the Stanley-Wilf conjecture, *J. Combinatorial Theory (Ser. A)* **107** (2004), 153–160.

[42] A. Méroueh, Lubell mass and induced partially ordered sets, ArXiv:150607056

[43] A. Methuku and D. Pálvőlgyi, Forbidden hypermatrices imply general bounds on induced forbidden subposet problems, *Combinatorics, Probability, and Computing* **26** (2017), 593-602

[44] A. Methuku and C. Tompkins, Exact forbidden subposet results using chain decompositions of the cycle, *Elect. J. Combinatorics* **22** (2015) #P4.29.

[45] B. Patkós, Supersaturation and stability for forbidden subposet problems, *J. Combinatorial Theory (Ser. A)* **136** (2015), 220–237.

[46] B. Patkós, Induced and non-induced forbidden subposet problems, *Elect. J. Combinatorics* **22** (2015) #P1.30.

[47] A. A. Razborov, Flag algebras, *Journal of Symbolic Logic* **72** (2007), 1239-1282.

[48] G. Sarkis, S. Shahriari, and PCURC, Diamond-free subsets in the linear lattices, *Order* **31** (2014), 421–43.

[49] E. Sperner, Ein Satz über Untermengen einer endlichen Menge, *Math. Z.* **27** (1928), 544–548.

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