CrossMark

The Journey of the Union-Closed Sets Conjecture

Henning Bruhn · Oliver Schaudt

Abstract We survey the state of the union-closed sets conjecture.

Keywords Union-closed · Frankl's conjecture · Graphs · Lattices

1 Introduction

One of the first mentions [6] of the union-closed sets conjecture calls it "a much-travelled conjecture". This is indeed so. Geographically it has spread from Europe to at least North America, Asia, Oceania and Australia. Mathematically it has ventured from its origins in extremal set theory to lattice and graph theory. In this survey we strive to trace its journey.

The main attraction of the conjecture is certainly its simple formulation. A family $\mathcal{A}$ of sets is *union-closed* if for every two member-sets $A, B \in \mathcal{A}$ also their union $A \cup B$ is contained in $\mathcal{A}$.
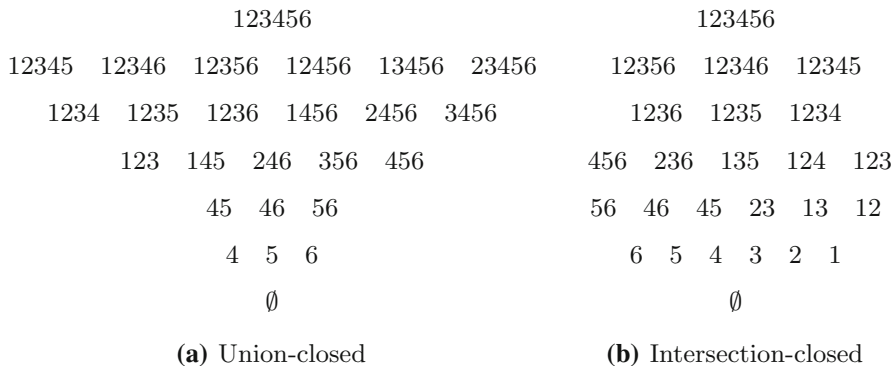
Union-closed sets conjecture *Any finite union-closed family of sets $\mathcal{A} \neq \{\emptyset\}$ has an element that is contained in at least half of the member-sets.*

An example of a union-closed family is given in Fig. 1a, where we have omitted commas and parentheses. There, one may count that the elements 1, 2, 3 appear each in only 12 of the 25 member-sets, which is less than half of the sets. Each of the other elements 4, 5, 6 however is contained in 16 sets, more than enough for the family to satisfy the conjecture. Power sets are other examples of union-closed families, and there the conjecture is tight: every element appears in exactly half of the member-sets.

H. Bruhn (✉)
Universität Ulm, Ulm, Germany
e-mail: henning.bruhn@uni-ulm.de

O. Schaudt
Institut für Informatik, Universität zu Köln, Weyertal 80, Cologne, Germany
e-mail: schaudto@uni-koeln.de

			123456					123456	
12345	12346	12356	12456	13456	23456	12356	12346	12345	

```
        123456                                    123456

12345  12346  12356  12456  13456  23456      12356  12346  12345

   1234  1235  1236  1456  2456  3456            1236  1235  1234

      123  145  246  356  456                456  236  135  124  123

         45  46  56                            56  46  45  23  13  12

            4   5   6                             6   5   4   3   2   1

                ∅                                      ∅

         (a) Union-closed                        (b) Intersection-closed
```

Fig. 1 A union-closed family and its complement

Despite its apparent simplicity the union-closed sets conjecture remains wide open. This is certainly not for lack of interest—there are about 50 articles dedicated to the conjecture, as well as several websites [48,68,69]. Due to this extensive research activity, we now know that the conjecture is satisfied for various union-closed families $\mathcal{A}$. For instance:

- if $\mathcal{A}$ has at most 12 elements or at most 50 member-sets;
- if the number n of member-sets is large compared to the number m of elements, that is, when $n \geq \frac{2}{3}2^m$;
- if n is small compared to m: when $n \leq 2m$ (where we need to assume that $\mathcal{A}$ is *separating*, that is, for any two elements there exists a member-set containing exactly one of them);
- if $\mathcal{A}$ contains one of a number of subconfigurations, such as a singleton-set;
- or if $\mathcal{A}$ has a particular structure, for instance, if $\mathcal{A}$ may be represented by a lower semimodular lattice, or by a subcubic graph.

We will discuss all these results, and give proper attributions, in the course of the article. All these partial results notwithstanding, we still seem to be far from a proof of the conjecture, and this is even the case for the obvious relaxation in which we settle for an element that appears in only, say, ≥ 1 % of the member-sets. The best result in this respect is an observation by Knill (slightly improved by Wójcik) that yields always an element of frequency at least $\frac{n-1}{\log_2 n}$.

In an article [7] of 1987, Peter Winkler[1] wrote "the 'union-closed sets conjecture' is well known indeed, except for (1) its origin and (2) its answer!" While the answer remains elusive, we can shed some light on its origins.

Most authors today attribute the conjecture to Peter Frankl, and following Frankl [26] date it to 1979. The sole exception are Balla et al. [9], who call it a "folklore conjecture" that "was well known by the mid-1970s". We cannot resolve this conflict of attribution, nor do we have the intention to do so. However, there is no

---

[1] Winkler informed us that the article was never intended to be published. Rather, this is the case of an informal letter ending up in print without Winkler even knowing.

doubt that Frankl did discover the conjecture (whether he was not the first is for others to decide) and that he played an instrumental role in popularising it. Consequently, we will sometimes speak of *Frankl's conjecture*.

In late 1979, Frankl [23] was working on traces of finite sets, a work that culminated in his article [24] of 1983. Motivated by the observation that it could be used to improve a number of bounds, Frankl formulated the conjecture when travelling from Paris to Montreal. On his way, Frankl told the conjecture to Ron Graham, who disseminated it widely. In about 1981, Dwight Duffus learnt about it, which then led to its first appearance in print: the proceedings of a workshop held in 1984 in Banff, edited by Rival [56], contain a short report of Duffus on a "problem of P. Frankl". The second mention is Stanley [64], which simply cites Rival.

The next time the conjecture appeared in print, it had apparently travelled with Franz Salzborn from Europe to Australia. An article of 1987 in the Australian Mathematical Society Gazette [6] reports on the Annual Meeting of the society during which Jamie Simpson publicised the conjecture. We may only speculate that this is how the conjecture arrived in Papua New Guinea, where Renaud and Sarvate went on to write the first published research articles about it [53,62,63] in 1989–1991. They were succeeded by Wójcik [70] in Poland and, in the USA, by Poonen [50], who wrote his influential article when he was an undergraduate. Many others followed in subsequent years.

In this survey, we aim to give a complete review of the literature on the conjecture. While we tried to track down every article with a substantial connection to the conjecture, we were not entirely successful as we could not obtain an unpublished manuscript of Salvi [60] that, as Wójcik [70] writes, apparently contains reformulations of the conjecture.

The focus of this survey is on the methods employed to attack the conjecture. Our treatment of the literature is therefore somewhat uneven. Whenever we can identify a technique that, to our eyes, seems interesting and potentially powerful we discuss it in greater detail.

2 Elementary Facts and Definitions

We quickly settle some notation and mention the most elementary facts. Let $\mathcal{A}$ be a family of sets. We call the set $U(\mathcal{A}) := \bigcup_{A \in \mathcal{A}} A$ of all the elements that appear in some member-set of $\mathcal{A}$ the *universe of* $\mathcal{A}$. If $\mathcal{A}$ is union-closed then taking the complements of all member-sets results in a family $\mathcal{D} = \{U(\mathcal{A}) \backslash A \colon A \in \mathcal{A}\}$ that is *intersection-closed*: if $C, D \in \mathcal{D}$ then also $C \cap D \in \mathcal{D}$.

The union-closed sets conjecture has the following equivalent form for intersection-closed families.

Intersection-closed sets conjecture *Any finite intersection-closed family of at least two sets has an element that is contained in at most half of the member-sets.*

Continuing with notation, we denote by

$$\mathcal{A}_x := \{A \in \mathcal{A} \colon x \in A\}.$$

the subfamily of member-sets containing any given element $x \in U(\mathcal{A})$. The cardinality $|\mathcal{A}_x|$ is the *frequency of* x in $\mathcal{A}$. We also introduce notation for the complement of $\mathcal{A}_x$:

$$\mathcal{A}_{\bar{x}} := \mathcal{A} \backslash \mathcal{A}_x = \{A \in \mathcal{A} \colon x \notin A\}.$$

We point out that, if $\mathcal{A}$ is union-closed, both $\mathcal{A}_x$ and $\mathcal{A}_{\bar{x}}$ are union-closed as well.

With this terminology, the union-closed sets conjecture states that in every (finite) union-closed family $\mathcal{A}$ there is an $x \in U(\mathcal{A})$ with $|\mathcal{A}_x| \geq \frac{1}{2}|\mathcal{A}|$. We will call such an element x *abundant*. When we consider an intersection-closed family $\mathcal{D}$, the intersection-closed sets conjecture asserts the existence of an element $y \in U(\mathcal{D})$ with $|\mathcal{D}_y| \leq \frac{1}{2}|\mathcal{D}|$. Such a y is *rare* in $\mathcal{D}$. (We realise that this leads to the slightly bizarre situation that an element with frequency $|\mathcal{A}_x| = \frac{1}{2}|\mathcal{A}|$ is at the same time abundant and rare.)

As Poonen [50] observed, the union-closed sets conjecture becomes false if the family is allowed to have infinitely many member-sets. Indeed, the union-closed family consisting of the sets $\{i, i+1, i+2, \ldots\}$ for every positive integer i has infinitely many member-sets but no element has infinite frequency. As a consequence, we will tacitly presuppose that every union-closed family considered in this survey has only finitely many member-sets.

Additionally, we will always require the universe to be finite as well. This is no restriction. If, for a union-closed family $\mathcal{A}$, the universe has infinite cardinality there will be infinitely many pairs of elements x and y in the universe of $\mathcal{A}$ that cannot be separated by $\mathcal{A}$, in the sense that $x \in A$ if and only if $y \in A$ for all $A \in \mathcal{A}$. In that case, we may simply delete y from all member-sets of $\mathcal{A}$. This results again in a union-closed family that satisfies the union-closed sets conjecture if and only if $\mathcal{A}$ does. Consequently, it suffices to prove the conjecture for *separating* families $\mathcal{A}$, those in which, for any two distinct elements $x, y \in U(\mathcal{A})$, there is an $A \in \mathcal{A}$ that contains exactly one of x, y. It is an easy observation that the universe of any (finite) separating family is finite.

We remark furthermore that, if necessary, we may always assume a union-closed family to include the empty set as a member. Adding $\emptyset$ will at most increase the number of sets, while obviously the frequency of any element stays the same. In the case of an intersection-closed family $\mathcal{D}$, it is no restriction to suppose that $\emptyset, U(\mathcal{D}) \in \mathcal{D}$. Indeed, adding $U(\mathcal{D})$ to $\mathcal{D}$ makes satisfying the intersection-closed sets conjecture only harder, while $\emptyset$ is always a member-set of $\mathcal{D}$ unless there is an element x appearing in every set of $\mathcal{D}$. In that case, deleting x from every member results in an intersection-closed family that satisfies the conjecture if and only if $\mathcal{D}$ does.

Given a family $\mathcal{S}$ of sets, the *union-closure* of $\mathcal{S}$ is the union-closed family $\mathcal{A}$ defined by

$$\mathcal{A} = \left\{ \bigcup_{S \in \mathcal{S}'} S \colon \mathcal{S}' \subseteq \mathcal{S} \right\}.$$

We may also say that $\mathcal{A}$ is *generated* by $\mathcal{S}$.

Every union-closed family $\mathcal{A}$ has a unique subset $\mathcal{B} \subseteq \mathcal{A}$ such that (a) $\mathcal{A}$ is the union-closure of $\mathcal{B}$ and (b) $\mathcal{B}$ is inclusionwise minimal with this property. Observe that $\mathcal{B}$ is simply the subfamily of non-empty sets $B \in \mathcal{A}$ with the property that if $B = X \cup Y$ for some $X, Y \in \mathcal{A}$, then $X = B$ or $Y = B$. The sets in $\mathcal{B}$ are the *basis sets* of $\mathcal{A}$. Observe that $\mathcal{A} \backslash \{B\}$ is union-closed for $B \in \mathcal{A}$ if and only if B is a basis set (or $B = \emptyset$).

Finally, for $i, n \in \mathbb{N}$ we use the notation $[n]$ to denote $\{1, \ldots, n\}$ and $[i, n]$ for the set $\{i, i + 1, \ldots, n\}$. We write 2^X for the power set of a set X. Any set of cardinality k is a *k-set*. For a set X and an element x, we often write $X + x$ for $X \cup \{x\}$ and $X - x$ for $X \backslash \{x\}$.

3 The Many Faces of the Conjecture

The union-closed sets conjecture has several equivalent reformulations that each highlight a different aspect. In this section we present three reformulations, one in terms of lattices, one in the language of graphs and the last again in terms of sets. That the same problem can be posed quite naturally in such different fields is a clear indication that Frankl's question is a very basic and fundamental one.

The reformulations also help us to gain confidence in the veracity of the conjecture. Indeed, each offers natural special cases such as semimodular lattices or subcubic graphs that would appear quite artificial in the other formulations. Proving the conjecture for such special cases then clearly adds evidence in support of the conjecture. Finally, each reformulation opens up new tools and techniques to attack the conjecture.

3.1 The Lattice Formulation

Already in its earliest mention [56] it is recognised that the union-closed sets conjecture, or rather its twin, the intersection-closed sets conjecture, has an equivalent formulation in terms of lattices. In fact, any intersection-closed[2] family together with inclusion forms a lattice.

We recall a minimum of lattice terminology. A *finite lattice* is a finite poset $(L, \leq)$ in which every pair $a, b \in L$ of elements has a unique greatest lower bound, denoted by $a \wedge b$ (the *meet*), and a unique smallest upper bound, denoted by $a \vee b$ (the *join*). All the lattices considered in this survey will be finite. The unique minimal element is denoted by 0, the unique maximal element is 1. A non-zero element $a \in L$ is *join-irreducible* if $a = b \vee c$ implies $a = b$ or $a = c$. We write $[a) := \{x \in L \colon x \geq a\}$. For more on lattices see, for instance, Grätzer [28].

Let us first see that an intersection-closed family $\mathcal{A}$ defines a lattice in a quite direct way. This is illustrated in Fig. 2, which shows the lattice corresponding to the family of Fig. 1b. As pointed out in the previous section, we may assume that $\mathcal{A}$ contains its universe $U(\mathcal{A})$. Then $(\mathcal{A}, \subseteq)$ is a lattice. Indeed, the unique greatest lower bound of any $A, B \in \mathcal{A}$ is $A \wedge B = A \cap B \in \mathcal{A}$, while $U(\mathcal{A}) \in \mathcal{A}$ guarantees that A and B

---

[2] Or union-closed family, for that matter. However, it seems customary in the lattice context to consider intersection-closed families.
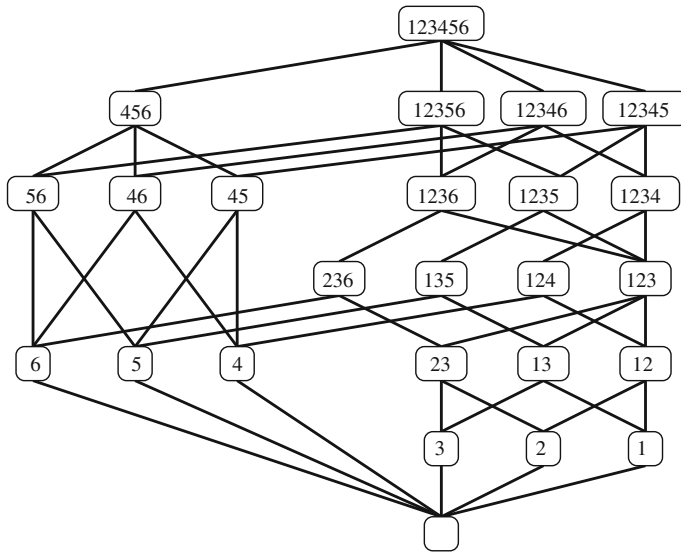
Fig. 2 The lattice of the set system in Fig. 1. The join-irreducible elements are precisely {1}, {2}, {3}, {4}, {5}, {6}

always have a minimal upper bound. Such a minimal upper bound is unique: if R and S are two upper bounds then also $R \cap S \in \mathcal{A}$ is an upper bound. Let us note that while $A \vee B$ always contains $A \cup B$, it is usually larger.

We now state the lattice formulation of Frankl's conjecture:

Conjecture 1 *Let L be a finite lattice with at least two elements. Then there is a join-irreducible element a with* $|[a]| \leq \frac{1}{2}|L|$.

Let us see why Conjecture 1 is equivalent to the intersection-closed sets conjecture. Let $\mathcal{A}$ be an intersection-closed family containing its universe and consider the lattice $(\mathcal{A}, \subseteq)$. Assume Conjecture 1 to hold, that is, there is a join-irreducible $J \in \mathcal{A}$ with $|[J]| \leq \frac{1}{2}|\mathcal{A}|$. Suppose that every element of J appears in some proper subset of J that is in $\mathcal{A}$: $\bigcup_{A \subset J} A = J$. Then, $\bigvee_{A \subset J} A \supseteq \bigcup_{A \subset J} A = J$, from which follows that $\bigvee_{A \subset J} A = J$, which is impossible as J is join-irreducible. Thus there is an $x \in J$ that does not lie in any proper subset of J.

Next, consider an $A \in \mathcal{A}$ containing x. Then $J \cap A$ is a subset of J containing x and therefore equal to J. In particular, $J \subseteq A$ and thus $A \in [J]$. Since $|[J]| \leq \frac{1}{2}|\mathcal{A}|$, it follows that x appears in at most half of the member-sets of $\mathcal{A}$.

For the other direction, consider a lattice L and associate to every $x \in L$ the set $S(x)$ of join-irreducible elements z with $z \leq x$. Then, for $x, y \in L$ we obtain that $S(x \wedge y) = S(x) \cap S(y)$, and thus the family $\mathcal{A} = \{S(x) : x \in L\}$ is intersection-closed. Moreover, $|\mathcal{A}| = |L|$.

Supposing that the intersection-closed sets conjecture holds, we obtain a join-irreducible $x \in L$ that is contained in at most half of the member-sets of $\mathcal{A}$. Then for any $y \geq x$, it follows that $x \in S(y)$ and thus $|[x]|$ is bounded by the number of member-sets of $\mathcal{A}$ containing x, which gives $|[x]| \leq \frac{1}{2}|L|$.

Theorem 2 *Conjecture 1 is equivalent to the union-closed sets conjecture.*

In view of this equivalence we will say that a lattice *satisfies Frankl's conjecture* if Conjecture 1 holds for it. To include the trivial case, we will extend this to any lattice on less than two elements.

What are the advantages of the lattice formulation? In some sense, Frankl's conjecture is stripped down to its bare essential parts: the elements have vanished and all that counts is the inclusion relation between the sets. Moreover, in comparison with the set formulation new special cases become natural—and attackable. We will review them next.

3.2 Lattice Results

The formulation of the lattice version resulted in a series of verified special cases of Frankl's conjecture. Already in Rival [56] it is mentioned, without proof, that the conjecture holds for distributive and geometric lattices. This was explicitly proved by Poonen [50], who also extended the latter case to complemented lattices.

Abe and Nakano [3] showed the conjecture for modular lattices, a case that includes distributive lattices. This, in turn, was generalised by Reinhold [52] to lower semimodular lattices. We present the proof here, as it seems to be the strongest result concerning lattice classes, and also because the proof is nice and succinct.

Let $x < y$ be two elements of a lattice. Then x is a *lower cover* of y if $x \leq z \leq y$ implies $x = z$ or $y = z$ for all elements z. A lattice L is *lower semimodular* if $a \wedge b$ is a lower cover of $a \in L$, whenever $b \in L$ is a lower cover of $a \vee b$.

Theorem 3 (Reinhold [52]) *Lower semimodular lattices satisfy Frankl's conjecture.*

Proof Let L be a lower semimodular lattice with $|L| \geq 2$. If the unique largest element $1 \in L$ is join-irreducible then Frankl's conjecture is trivially satisfied. If not, we may pick a lower cover $b \in L$ of 1, and a join-irreducible $a \in L$ with $a \nleq b$. Then $1 = a \vee b$.

We claim that the function $[a) \to L \backslash [a)$, $x \mapsto x \wedge b$ is an injection, which then finishes the proof. So, suppose that there are two distinct $x, y \in [a)$ with $x \wedge b = y \wedge b$. As either $x \wedge y < x$ or $x \wedge y < y$, we may assume the former. This implies

$$x \wedge b = x \wedge y \wedge b \leq x \wedge y < x. \tag{1}$$

Now, as L is lower semimodular, and as b is a lower cover of $1 = x \vee b$, we obtain that $x \wedge b$ is a lower cover of x. Thus, $x \wedge b = x \wedge y$ by (1) and therefore

$$a \leq x \wedge y = x \wedge b \leq b,$$

which contradicts our choice of $a \nleq b$.   □

Theorem 3 was also independently proved by Herrmann and Langsdorf [29] and by Abe and Nakano [4]. In the latter article, the conjecture is also verified for a superclass, lower quasi-semimodular lattices.

If there are lower semimodular lattices there are clearly *upper semimodular* ones as well. However, this class seems to be much harder with respect to Frankl's conjecture. Already in Rival [56] it is mentioned, without proof, that *geometric lattices* satisfy the conjecture. A proper proof was later given by Poonen [50]. A lattice is geometric, and then upper semimodular, if it may be represented as the lattice of flats of a matroid. Abe [1] treats another subclass, the so called *strong* upper semimodular lattices. Czédli and Schmidt [15] show the conjecture for upper semimodular lattices L that are large, in the sense that $|L| > \frac{5}{8}2^m$ where m is the number of join-irreducible elements; they also consider planar upper semimodular lattices.

Let us mention that it is an easy consequence of the lattice formulation that, for any lattice L, Frankl's conjecture holds for L or for its dual L^*, or both. (The dual lattice is obtained by reversing the order.) Duffus and Sands [18] and Abe [2] derive stronger assertions for special classes of lattices.

We close this section with a wonderful application of Reinhold's theorem that was indicated to us by one of the anonymous referees. The application concerns *graph-generated* intersection-closed families. Let G be a fixed graph. For every set $X \subseteq V(G)$ we write E_X for the set of edges of G that have both their endvertices in X. Then $\{E_X : X \subseteq V(G)\}$ is intersection-closed.

Theorem 4 (Knill [37]) *Given a graph $G = (V, E)$ with at least one edge, the intersection-closed family $\{E_X : X \subseteq V\}$ satisfies the intersection-closed sets conjecture.*

This result is also part of Knill's PhD thesis [36]. The theorem was later restated as a conjecture by El-Zahar [19], and, as a response to El-Zahar's paper, reproved by Llano et al. [41].

As $L = \{E_X : X \subseteq V(G)\}$ is intersection-closed, it is a lattice with respect to $\subseteq$. We show that L is lower semimodular. Thus, Knill's theorem becomes a consequence of Theorem 3.

We call $X \subseteq V(G)$ *proper* if $E_X \neq E_{X'}$ for any $X' \subsetneq X$. Note that $L = \{E_X : X \subseteq V(G) \text{ and } X \text{ is proper}\}$, and so we may restrict our attention to proper vertex sets. Let $X, Y \subseteq V(G)$ be proper. First we note that

$$E_X \wedge E_Y = E_X \cap E_Y = E_{X \cap Y} \quad \text{and} \quad E_X \vee E_Y = E_{X \cup Y}.$$

Next we observe that E_X is a lower cover of E_Y if and only if

$$Y = X + y_1 \quad \text{or} \quad E_Y = E_X + y_1 y_2 \quad \text{for some } y_1, y_2 \in Y \backslash X.$$

Indeed, let E_X be a lower cover of E_Y and consider an edge $y_1 y_2 \in E_Y \backslash E_X$. Then, $E_X \subsetneq E_{X \cup \{y_1, y_2\}} \subseteq E_Y$ and thus $Y = X \cup \{y_1, y_2\}$. Now, if one of y_1, y_2 say, is contained in X we have $Y = X + y_1$ and we are in the first case. If $y_1, y_2 \notin X$ then neither of y_1, y_2 may have a neighbour in X as otherwise E_X would be a proper subset of E_{X+y_1} or of E_{X+y_2}. The other direction is obvious.

So, assume that for proper $A, B \subseteq V(G)$, the set E_B is a lower cover of $E_A \vee E_B$. Then there are $a_1, a_2 \in A \backslash B$ so that either $A \cup B = B + a_1$ or $E_{A \cup B} = E_B + a_1 a_2$.

If $A \cup B = B + a_1$ then $A = (A \cap B) + a_1$, and $E_{A \cap B}$ is a lower cover of E_A. In the other case, when $E_{A \cup B} = E_B + a_1 a_2$ we get

$$E_A = E_A \cap E_{A \cup B} = (E_A \cap E_B) + a_1 a_2 = E_{A \cap B} + a_1 a_2,$$

and again $E_{A \cap B}$ is a lower cover of E_A. Thus, L is lower semimodular, and Knill's theorem is proved.

El-Zahar [19] observed that, when Knill's theorem is generalised to hypergraphs, it becomes yet another reformulation of the union-closed sets conjecture.

3.3 The Graph Formulation

A more recent reformulation of the union-closed sets conjecture is stated in terms of maximal stable sets of bipartite graphs. A *stable set* of a graph G is a vertex subset so that no two of its vertices are adjacent. A stable set is called *maximal* if no further vertex of G can be added without violating the stable set condition. We refer to Diestel [16] for general terminology and notions on graphs.

The graph formulation of the union-closed sets conjecture is as follows:

Conjecture 5 *Any bipartite graph with at least one edge contains in each of its bipartition classes a vertex that lies in at most half of the maximal stable sets.*

The conjecture was proposed by Bruhn et al. [11], who also proved the equivalence to Frankl's conjecture. In analogy to the intersection-closed sets conjecture, let us call a vertex *rare* if it is contained in at most half of the maximal stable sets. Note that for every edge uv of a bipartite graph, always one of u and v is rare. Indeed, this follows directly from the fact that no stable set may contain both u and v. Hence, in a hypothetical counterexample to Conjecture 5, one bipartition class of the graph contains only rare vertices, while no vertex in the other class is rare.

We sketch why Conjecture 5 and the intersection-closed sets conjecture are equivalent.

**Theorem 6** [11] *Conjecture 5 holds if and only if the union-closed sets conjecture is true.*

Proof To prove equivalence to the intersection-closed sets conjecture, let us first consider a bipartite graph G with bipartition classes X, Y. By symmetry it is enough to find a rare vertex in X. Let $\mathcal{A}$ be the set of maximal stable sets of G. It is straightforward to check that the traces of maximal stable sets in X, the set $\{A \cap X : A \in \mathcal{A}\}$, is intersection-closed. Thus, if the intersection-closed sets conjecture is true, there must be a rare element x of $\{A \cap X : A \in \mathcal{A}\}$, which then is a rare vertex of G.

For the converse direction, let an intersection-closed family $\mathcal{A}$ be given. We may assume that $\mathcal{A}$ contains its universe U. We define a bipartite graph $G = (V, E)$ on $V = \mathcal{A} \cup U$ with edge set $E = \{Sx : S \in \mathcal{A}, x \in U, x \in S\}$. That is, G is the incidence graph of $\mathcal{A}$. See Fig. 3 for an illustration.

Then, if $\mathcal{B}$ denotes the set of maximal stable sets of G, it follows that $\mathcal{A} = \{B \cap U : B \in \mathcal{B}\}$. Thus, if x is a rare vertex of G in U, then x is a rare element of $\mathcal{A}$. This completes the proof. □
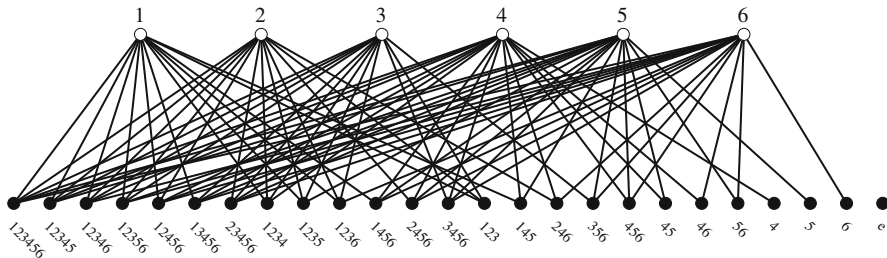
Fig. 3 The incidence graph of the intersection-closed family shown in Fig. 1
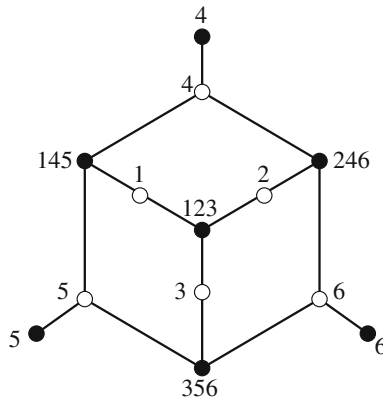
Fig. 4 A more succinct representation

As for the lattice fromulation, we will say that a bipartite graph *satisfies Frankl's conjecture* if the graph is not a counterexample to Conjecture 5, or if it is edgeless.

Figure 3 shows the graph representation of intersection-closed family in Fig. 1. We have to admit that it does not appear very appealing, as listing the family seems much simpler. Nonetheless, the graph formulation allows for a very compact representation of Frankl's conjecture. This is exemplified by the graph in Fig. 4 that encodes the same family as the graph in Fig. 3. We arrive at this graph by iteratively deleting any vertex v whose neighbourhood is equal to the union of neighbourhoods of some other vertices. It is easy to check that the resulting graph with v deleted satisfies the conjecture only if the original graph does, see also [11].

3.4 Graph Results

The literature on graphs provides a rich selection of natural graph classes, even bipartite ones, that may now serve as test cases for Frankl's conjecture. So far, the conjecture has been verified for chordal bipartite, subcubic, series-parallel [11] and, in an approximate version, random bipartite graphs [12]. We present some of these results here.

A bipartite graph is said to be *chordal bipartite* if deleting vertices from the graph can never result in a chordless cycle of length ≥ 6.

Theorem 7 [11] *Chordal bipartite graphs satisfy Frankl's conjecture.*

The proof rests on the local structure of chordal bipartite graphs. This is a general strategy that we will discuss in more detail in Sect. 5. The main tool here is the following lemma, where we denote by $N^2(x)$ the neighbours of the neighbours of a vertex x (including x).

Lemma 8 [11] *Let x, y be two adjacent vertices of a bipartite graph with $N^2(x) \subseteq N(y)$. Then y is rare.*

Proof Let $\mathcal{A}$ denote the maximal stable sets of the chordal bipartite graph G, and consider $A \in \mathcal{A}_y$, that is, a maximal stable set containing y. Since $y \in A$, no neighbour of y may be in A and hence $N^2(x) \cap A = \emptyset$ as $N^2(x) \subseteq N(y)$. Therefore, no vertex in $N(x)$ is adjacent with a vertex in A, which implies $N(x) \subseteq A$.

We now construct an injective mapping $\mathcal{A}_y \rightarrow \mathcal{A}_x$: given a set $A \in \mathcal{A}_y$, first remove all members of $N(x)$ from A and then fill up the resulting set to a maximal stable set with vertices from $N^2(x)$. Finally, since x is adjacent to y, we have $\mathcal{A}_x \subseteq \mathcal{A}_{\overline{y}}$. Altogether, there is an injection $\mathcal{A}_y \rightarrow \mathcal{A}_{\overline{y}}$, which means that y is rare.   □

To finish the proof of Theorem 7 it now suffices to observe that a type of vertex known as a *weakly simplicial* vertex satisfies the conditions of the lemma. That such a vertex always exists in each bipartition class is known from the literature on chordal bipartite graphs. For details see [11].

Using results of Vaughan on 3-sets and Knill's graph generated families (discussed in Sects. 5, 3.2 respectively), we obtain Frankl's conjecture for another natural graph class. Recall that a graph is *subcubic* if every vertex has degree at most three.

Theorem 9 [11] *Every subcubic bipartite graph satisfies Frankl's conjecture.*

The third class of graphs we treat are random bipartite graphs, where we can only prove a slight weakening of Frankl's conjecture. A *random bipartite graph* is a graph on bipartition classes of cardinalities m and n, where any two vertices from different classes are independently joined by an edge with probability p.

For $\delta > 0$, let us say that a bipartite graph *satisfies Frankl's conjecture up to δ* if each of its two bipartition classes has a vertex for which the number of maximal stable sets containing it is at most $\frac{1}{2} + \delta$ times the total number of maximal stable sets. We say that *almost every* random bipartite graph has property P if for every $\varepsilon > 0$ there is an N such that, whenever $m + n \geq N$, the probability that a random bipartite graph on $m + n$ vertices has P is at least $1 - \varepsilon$.

Theorem 10 [12] *Let $p \in (0, 1)$ be a fixed edge-probability. For every $\delta > 0$, almost every random bipartite graph satisfies Frankl's conjecture up to δ.*

The main tool in the proof is the averaging approach detailed in Sect. 6.

3.5 The Salzborn Formulation

Returning to the sets point of view, let us present a surprising reformulation of the conjecture that Wójcik [71] attributes to Salzborn [61]. Recall that a union-closed

family $\mathcal{A}$ is separating if for any two elements of its universe there is a member-set that contains exactly one of the two. It is easy to check that $\mathcal{A}$ needs to have at least $|U(\mathcal{A})|$ non-empty sets to separate all elements of its universe. Thus, if $\emptyset \in \mathcal{A}$ then $\mathcal{A}$ will have at least $|U(\mathcal{A})| + 1$ member-sets. It turns out that the families with this minimum number of member-sets have a surprisingly rich structure.

Let us call a union-closed family $\mathcal{N}$ *normalised* if it holds that $\emptyset \in \mathcal{N}$, $\mathcal{N}$ is separating and $|U(\mathcal{N})| = |\mathcal{N}| - 1$. The following conjecture may be found in Wójcik [71], or, with less details, in Salzborn [61].

Conjecture 11 (Salzborn [61]) *Any normalised family $\mathcal{N} \neq \{\emptyset\}$ contains a basis set B of size $|B| \geq \frac{1}{2}|\mathcal{N}|$.*

Following Wójcik [71], we outline why Salzborn's conjecture implies the union-closed sets conjecture. Consider a union-closed family $\mathcal{A}$ that we may assume to contain $\emptyset$ as a member-set. We define

$$\mathcal{A}_{\nsubseteq X} := \{A \in \mathcal{A} \colon A \nsubseteq X\} \quad \text{and} \quad \mathcal{A}^* := \left\{ \mathcal{A}_{\nsubseteq X} : X \in \mathcal{A} \right\}.$$

It is easy to check that $\mathcal{A}^*$ is union-closed and separating. We note that $X \subseteq Y$ if and only if $\mathcal{A}_{\nsubseteq X} \supseteq \mathcal{A}_{\nsubseteq Y}$ for any $X, Y \in \mathcal{A}$. This has several consequences. Firstly, $\mathcal{A}_{\nsubseteq X} \neq \mathcal{A}_{\nsubseteq Y}$ if $X \neq Y$, which implies that $|\mathcal{A}| = |\mathcal{A}^*|$. Secondly, $U(\mathcal{A}^*) = \mathcal{A}_{\nsubseteq \emptyset} = \mathcal{A} \setminus \{\emptyset\}$. Finally, we remark that $\mathcal{A}^*$ has the dual lattice structure of $\mathcal{A}$.

To summarise, $\mathcal{A}^*$ is normalised and has the same number of members as $\mathcal{A}$. Next, we consider the basis sets of $\mathcal{A}^*$.

$$\text{Every basis set of } \mathcal{A}^* \text{ is of the form } \mathcal{A}_x \text{ for some,} \quad x \in U(\mathcal{A}). \tag{2}$$

Indeed, consider a basis set $\mathcal{A}_{\nsubseteq X}$ of $\mathcal{A}^*$, and observe that $\mathcal{A}_{\nsubseteq X} = \bigcup_{y \in U(\mathcal{A}) \setminus X} \mathcal{A}_y$. Pick a smallest set $S \subseteq U(\mathcal{A}) \setminus X$ so that still $\mathcal{A}_{\nsubseteq X} = \bigcup_{y \in S} \mathcal{A}_y$ and consider a bipartition $S_1 \cup S_2 = S$. Since $\mathcal{A}_y = \mathcal{A}_{\nsubseteq U(\mathcal{A}_{\bar{y}})}$, both $\bigcup_{y \in S_1} \mathcal{A}_y$ and $\bigcup_{y \in S_2} \mathcal{A}_y$ are members of $\mathcal{A}^*$. Since $\mathcal{A}_{\nsubseteq X} = \bigcup_{y \in S_1} \mathcal{A}_y \cup \bigcup_{y \in S_2} \mathcal{A}_y$, $\mathcal{A}_{\nsubseteq X}$ is the union of two member-sets of $\mathcal{A}^*$. As $\mathcal{A}_{\nsubseteq X}$ is a basis set that implies that already $\mathcal{A}_{\nsubseteq X} = \bigcup_{y \in S_i} \mathcal{A}_y$ for $i = 1$ or $i = 2$, which by the minimality of S forces $S = S_i$. Therefore, S has to contain a unique element x, that is, $\mathcal{A}_{\nsubseteq X} = \mathcal{A}_x$.

Assume now Conjecture 11 to hold. Then the normalised family $\mathcal{A}^*$ contains a basis set B^* with

$$|B^*| \geq \tfrac{1}{2}|\mathcal{A}^*| = \tfrac{1}{2}|\mathcal{A}|.$$

As $B^* = \mathcal{A}_x$ for some $x \in U(\mathcal{A})$ by (2) we deduce that $\mathcal{A}$ satisfies the union-closed sets conjecture. We therefore have proved one direction of:

Theorem 12 (Salzborn [61]) *Conjecture 11 is equivalent to the union-closed sets conjecture.*

We omit the proof of the other direction, which may be found in Wójcik [71].

Why do we find the Salzborn reformulation surprising? At first glance, normalised families seem to be very restricted and in some sense this is true. For instance, the statement of the union-closed sets conjecture is almost trivial for them, see Theorem 23. From a lattice point of view, however, normalised families turn out to be as general as union-closed families. We have already remarked that $\mathcal{A}^*$ has the dual lattice structure of $\mathcal{A}$, which directly implies that every lattice type of a union-closed family is realisable as a normalised family.

We know only one application of the Salzborn formulation: Wójcik [71] uses it to obtain a non-trivial lower bound on the maximum frequency of an element in a union-closed family; see the next section.

The family $\mathcal{A}^*$ also appears in Johnson and Vaughan [33], although defined in a slightly different way. In order to obtain a duality result, Johnson and Vaughan associate to any union-closed family $\mathcal{A}$ the dual family $\mathcal{A}^*$ and then observe that the union-closed sets conjecture is satisfied for at least one of $\mathcal{A}$ and $\mathcal{A}^*$. We note that the analogous results in the lattice formulation and in the graph formulation are almost trivial: for lattices this amounts to considering the dual lattice, and for graphs it reduces to the observation that no stable set may contain both endvertices of an edge.

The majority of the results on the union-closed sets conjecture are with respect to the original set formulation. In the remainder of this article we stick to this formulation as well. However, a good part of the discussed techniques has a more or less direct analogue in the other formulations.

4 Obstacles to a Proof

There are many results on special cases of the conjecture. Amazingly, if we consider an arbitrary union-closed family, without any special structure or information on the number of elements, (almost) the best result we have seems to be a simple observation due to Knill:

Theorem 13 (Knill [37]) *Any union-closed family $\mathcal{A}$ on n member-sets has an element of frequency at least $\frac{n-1}{\log_2(n)}$.*

Proof We may assume that $\emptyset \in \mathcal{A}$. Let us choose $S \subseteq U(\mathcal{A})$ minimal such that every non-empty set of $\mathcal{A}$ intersects S. Then for every $x \in S$ there is a $A \in \mathcal{A}$ with $A \cap S = \{x\}$; otherwise $S - x$ would still meet every non-empty $A \in \mathcal{A}$, which contradicts the minimality of S. As $\mathcal{A}$ is union-closed it follows that $\{A \cap S \colon A \in \mathcal{A}\} = 2^S$. Hence $n \geq 2^{|S|}$ and so $|S| \leq \log_2(n)$. As every of the $n - 1$ non-empty member-sets of $\mathcal{A}$ intersects S, there is an element in S that belongs to at least $(n - 1)/\log_2(n)$ many member-sets of $\mathcal{A}$.   □

Wójcik [71] improved the bound to $\frac{2.4n}{\log_2 n}$ for large n. His proof is not trivial, but the result is still far from Frankl's conjecture.

Here are two observations that could be interpreted as signs that the conjecture is, after all, perhaps not as hard as thought: normally the most frequent element appears more often than needed, and there are several abundant elements. Indeed, the powerful averaging technique discussed in Sect. 6 builds solely on these facts.

These observations are due to Poonen, who also found exceptions to them. Power sets are an obvious example for families in which the maximum frequency is exactly half the size of the family. Poonen conjectured that, among separating families, these are the only ones.

Conjecture 14 (Poonen [50]) *Let $\mathcal{A}$ be a separating union-closed family. Unless $\mathcal{A}$ is a power set, it contains an element that appears in strictly more than half of the member-sets of $\mathcal{A}$.*

A similar conjecture was offered by Renaud [53]. Moreover, Poonen described families with a unique abundant element and again conjectured that these are the only ones:

Conjecture 15 (Poonen [50]) *Let $\mathcal{A}$ be a separating union-closed family on universe U. If $\mathcal{A}$ contains a unique abundant element a then*

$$\mathcal{A} = \{\emptyset\} \cup \{B + a \colon B \subseteq U - a\}.$$

If these conjectures are to be believed, then there is a bit of a margin when attacking the union-closed sets conjecture. So, why then has the conjecture withstood more than 20 years of proof attempts?

The obvious first approach is to try an induction, for instance on the number of member-sets. If, given a union-closed family, we could delete one (or two) basis sets so that the maximum frequency drops then, by induction, the original family would satisfy the conjecture, too. Unfortunately, this is not always possible: in a power set of sufficient size, deleting one or two basis sets will never reduce the maximum frequency.

So, naive induction will not succeed. Often, induction can only be made to work if the hypothesis is strengthened, usually by exploiting some structural insight. However, we feel that we are lacking in just that. We do not know what the extremal families look like, those that have minimal maximum frequency among all union-closed families of a given size. So far, there are not even any good candidates. We will continue this discussion in Sect. 8.

A second reason why the conjecture has resisted so long lies in the weakness of the techniques at our disposal. Let us briefly review the main techniques used to prove that a given family satisfies the conjecture: *injections*, *local configurations* and *averaging*. In averaging we try to show that the average frequency is large enough so that some element must be abundant. Averaging is very powerful but has the drawback that there are families for which the average is simply too low for the method to work. We discuss averaging and its limits in Sect. 6. For the local configurations method one strives to identify small families so that any large union-closed family containing the small one will automatically satisfy the conjecture. Unfortunately, given what we know at the moment it seems doubtful that we will be able to show that any union-closed family always contains such a local configuration. We will have a closer look at local configurations in the next section.

That leaves injections, the simplest of the three techniques. For an almost trivial example, consider the case when a union-closed family $\mathcal{A}$ contains a singleton, that is, there is an element x so that $\{x\} \in \mathcal{A}$. Then

$$\mathcal{A}_{\overline{x}} \to \mathcal{A}_x, \quad A \mapsto A + x$$

defines an injection, which clearly implies that $2|\mathcal{A}_x| \geq |\mathcal{A}_x| + |\mathcal{A}_{\overline{x}}| = |\mathcal{A}|$. Consequently, x is abundant. In fact, we have used this method already twice: once for lower semimodular lattices and then for chordal bipartite graphs. The main problem with the injection method is that we need to first identify an element that is likely to be abundant.

Sarvate and Renaud [62] were probably the first to observe (in print) that a singleton is always abundant. In a similar way, one of the two elements of any 2-set is abundant. The pattern, however, breaks with 3-sets. Sarvate and Renaud [63] describe a family with a unique smallest member-set of 3 elements, none of which is abundant. Poonen [50] constructs a similar family, a generalisation of which we present here:

For each $k \geq 3$ we define a union-closed family $\mathcal{A}^k$ with the property that $[k]$ is the unique smallest set, but no element of $[k]$ is abundant. For this, we use Poonen's notation $\mathcal{A} \uplus \mathcal{B}$ for two set families $\mathcal{A}$ and $\mathcal{B}$ to denote the family

$$\mathcal{A} \uplus \mathcal{B} := \{S \cup T : S \in \mathcal{A}, \, T \in \mathcal{B}\}.$$

Now let

$$\mathcal{A}^k = \{[k]\} \cup \bigcup_{i=1}^{k}(\{\emptyset, \{i\}, [k]\} \uplus \mathcal{B}^i) \cup (2^{[k]} \uplus [k+1, 3k]),$$

where

$$\mathcal{B}^i = \{[k+1, 3k]\backslash\{2i+2\}, [k+1, 3k]\backslash\{2i+3\}\} \quad \text{for every } i \in [k].$$

Note that the set $[k]$ is the unique smallest set in $\mathcal{A}$. In total, $\mathcal{A}^k$ contains $1 + 6k + 2^k$ many sets, but every $i \in [k]$ is contained in exactly $1 + (2k + 2) + 2^{k-1}$ sets of $\mathcal{A}$. Therefore, no element of $[k]$ is abundant.

Poonen's family highlights one of the major obstacles on the way to a proof of the union-closed sets conjecture: we do not know where to expect an abundant element. However, there are special cases where this is known. We treat these cases next.

5 Local Configurations

Sarvate and Renaud [62] observed that any singleton in a union-closed family is abundant, and of the two elements of a 2-set at least one is abundant. This motivates the search for good *local configurations*: a family $\mathcal{L}$ on few elements so that any union-closed family $\mathcal{A}$ containing $\mathcal{L}$ has an abundant element among the elements of $\mathcal{L}$. Poonen [50] gives a complete characterisation of such families:

Theorem 16 (Poonen [50]) *Let $\mathcal{L}$ be a union-closed family with universe $[k]$. The following statements are equivalent:*

(i) *Every union-closed family $\mathcal{A}$ containing $\mathcal{L}$ satisfies the union-closed sets conjecture. In particular, $\mathcal{A}$ has an abundant element in $[k]$.*

(ii) *There are reals $c_1, c_2, \ldots, c_k \geq 0$ with $\sum_{i=1}^{k} c_i = 1$ such that for every union-closed family $\mathcal{K} \subseteq 2^{[k]}$ with $\mathcal{K} = \mathcal{L} \uplus \mathcal{K}$ it holds that*

$$\sum_{i=1}^{k} c_i |\mathcal{K}_i| \geq \tfrac{1}{2}|\mathcal{K}|.$$

We stress that (ii) is indeed a local condition: for fixed k there are only finitely many such families $\mathcal{K}$. As an application of his theorem, Poonen showed that the union-closed family consisting of a 4-set together with any three distinct 3-subsets satisfies the conditions of his theorem. This was later generalised by Vaughan [67] to three distinct 3-sets with a non-empty common intersection. As mentioned in Sect. 3.4, Vaughan's result is used to prove Frankl's conjecture for subcubic bipartite graphs.

A union-closed family $\mathcal{L}$ as in Theorem 16 is called *Frankl-complete* by Vaughan [66], *FC* for short. Several FC-families are listed in [66], for example a 5-set together with all its 4-subsets or a 6-set with all 5-subsets and eight 4-subsets. The list was later extended by Morris [45], who, in particular, completely characterised the FC-families on at most 5 elements.

To study FC-families in a more quantitative way, Morris [45] introduced the function $FC(k, m)$ defined as the smallest r for which the set of every r of the k-sets in $[m]$ generates an FC-family. He showed that $\lfloor \frac{m}{2} \rfloor + 1 \leq FC(3, m)$, while Vaughan [67] gave an upper bound of $FC(3, m) \leq \frac{2m}{3}$. A proof of Morris' conjecture that $FC(3, m) = \lfloor \frac{m}{2} \rfloor + 1$ was announced by Vaughan [65], but has apparently never been published.

Marić et al. [43] verified some known FC-families and found a new one using the automatic proof assistant Isabelle/HOL. For this, they formalised the condition of FC-families to enable a computer search. As a result, we know now that all families containing four 3-subsets of a 7-set are FC-families.

5.1 Small Finite Families

The union-closed sets conjecture has been verified for families on few member-sets or few elements. The current best results use local configurations to reduce the number of special cases substantially.

With respect to the size of the universe, the conjecture has to-date been verified up to $m = 12$:

Theorem 17 (Živković and Vučković [72]) *The union-closed sets conjecture holds for union-closed families on at most 12 elements.*

The following result, that has not been improved upon in the last 20 years, allows to leverage bounds on the universe size to bounds on the number of member-sets:

Lemma 18 (Faro [22]) *Under the assumption that the union-closed sets conjecture fails, let m denote the minimum cardinality of $|U(\mathcal{A})|$ taken over all counterexamples $\mathcal{A}$ to the union-closed sets conjecture. Then any counterexample has at least $4m - 1$ member-sets.*

The lemma was later rediscovered by Roberts and Simpson [58]. Together with Theorem 17 we obtain:

Corollary 19 *The union-closed sets conjecture holds for union-closed families with at most 50 sets.*

Various authors verified the conjecture for small values of n and m, where as usual n is the number of member-sets and m the size of the universe. The first were Sarvate and Renaud [62] who treated a close variant that excludes the empty set. In a first paper they covered all cases up to $n \leq 11$; in Sarvate and Renaud [63] the case analysis was pushed up to $n \leq 19$. Using his Theorem 16, Poonen improved the bounds to $m \leq 7$ and $n \leq 28$. This was followed by Faro [22], who settled the union-closed sets conjecture for $m \leq 9$ and $n \leq 36$. For this, he investigated several necessary conditions on a minimal counterexample, among them Lemma 18 above. Roberts [57] shows the conjecture up to $n \leq 40$.

Using the list of known FC-families, Morris [45] proved the union-closed sets conjecture for families with $m \leq 9$ and $n \leq 36$, apparently unaware of the older result by Faro [22]. Nevertheless, there is merit in Morris' proof as it showcases how FC-families may be used to substantially reduce the number of cases. This method is at the heart of all subsequent work in this direction.

In order to prove the conjecture for $m \leq 10$, Marković [44] imitated the method of Theorem 16: he assigns non-negative weights to the elements of $\mathcal{A}$ and extends this to the member-sets of $\mathcal{A}$. He then observes that a total weight of the member-sets of at least $\frac{1}{2}n$ times the weight of the universe is sufficient for the union-closed sets conjecture. As a by-product of this method, Marković discovered a number of new FC-families.

Bošnjak and Marković [10] improve upon [44] by developing more general local configurations that allow them to verify the conjecture up to $m = 11$. With a very similar method and the use of a computer, Živković and Vučković [72] pushed this to $m \leq 12$.

6 Averaging

Obviously, a union-closed family $\mathcal{A}$ has an element of frequency $\geq \frac{1}{2}|\mathcal{A}|$ if the *average frequency* is at least $\frac{1}{2}|\mathcal{A}|$. In other words, if

$$\frac{1}{|U(\mathcal{A})|} \cdot \sum_{u \in U(\mathcal{A})} |\mathcal{A}_u| \geq \frac{1}{2}|\mathcal{A}|, \tag{3}$$

then $\mathcal{A}$ satisfies the union-closed sets conjecture.

So far, not much is gained. Calculating $\sum_{u \in U(\mathcal{A})} |\mathcal{A}_u|$ directly is clearly out of question, as this would presuppose knowledge about the individual frequencies $|\mathcal{A}_u|$. Fortunately, this is not necessary, as the sum of frequencies can be determined indirectly with a simple double-counting argument:

$$\sum_{u \in U(\mathcal{A})} |\mathcal{A}_u| = \sum_{A \in \mathcal{A}} |A|. \tag{4}$$

This identity is the heart of the averaging method. The total set size is usually much easier to control, and in some cases may be estimated quite well.

Combining (3) and (4), a condition equivalent to (3) is that

$$\frac{1}{|\mathcal{A}|} \cdot \sum_{A \in \mathcal{A}} |A| \geq \frac{1}{2} |U(\mathcal{A})|.$$

That is, if the *average set size* of $\mathcal{A}$ is at least half the size of the universe then $\mathcal{A}$ again satisfies the union-closed sets conjecture.

As discussed in Sect. 4, it is not obvious where to look for an abundant element. The averaging method has the clear advantage that it simply sidesteps this obstacle. In this section we describe how both (3) and (4) lead to some of the strongest results on the union-closed sets conjecture.

6.1 Large Families

In a clearly overlooked paper, Nishimura and Takahashi [46] prove for the first time that the union-closed sets conjecture always holds for large families. Their proof uses the average set size argument: it is shown that the average set size is greater than $\frac{m}{2}$, which implies that there is an abundant element.

Theorem 20 (Nishimura and Takahashi [46]) *Let $\mathcal{A}$ be a union-closed family of more than $2^m - \frac{1}{2}\sqrt{2^m}$ member-sets on a universe of size m. Then $\mathcal{A}$ satisfies the union-closed sets conjecture.*

Proof Suppose there is a set $S \subseteq U(\mathcal{A})$ with $S \notin \mathcal{A}$ but $|S| \geq \frac{m}{2}$. Then for any subset $R \subseteq S$ with $R \in \mathcal{A}$ it holds that $S \backslash R \notin \mathcal{A}$. Thus, at least half of the subsets of S are missing in $\mathcal{A}$. This gives $|\mathcal{A}| \leq 2^m - \frac{1}{2} \cdot 2^{\frac{m}{2}}$, a contradiction. Hence, every set $S \subseteq U(\mathcal{A})$ of size at least $\frac{m}{2}$ is contained in $\mathcal{A}$. This means that the average set size is at least $\frac{m}{2}$, finishing the proof.   □

Czédli [13] employed some involved lattice-theoretic arguments to push the bound from $2^m - \frac{1}{2}\sqrt{2^m}$ to $2^m - \sqrt{2^m}$. A weaker result than Nishimura and Takahashi's was proved by Gao and Yu [27]. Recently, a serious improvement of the above bound was given by Balla et al. [9], which we present in Sect. 6.4.

6.2 Bounds on the Average

Averaging does not always work. It is easy to construct union closed families with an average frequency and average set size that is too low to deduce the union-closed sets conjecture. Reimer [51] gave a bound on the average set size that is in some respect best possible.

Theorem 21 (Reimer [51]) *Let $\mathcal{A}$ be a union-closed family on n sets. Then*

$$\frac{1}{n} \cdot \sum_{A \in \mathcal{A}} |A| \geq \frac{\log_2 n}{2}. \tag{5}$$

The result is too weak for Frankl's conjecture as usually $\log_2(n) < m$. In terms of the average frequency, Reimer's bound reads as

$$\frac{1}{m} \cdot \sum_{u \in U(\mathcal{A})} |\mathcal{A}_u| \geq \frac{\log_2 n}{m} \cdot \frac{n}{2}. \tag{6}$$

We discuss the beautiful proof of Theorem 21 in Sect. 6.4.

We now focus on separating union-closed families, where for every two elements there is a set containing exactly one of them. As explained in Sect. 2, for the purpose of the union-closed sets conjecture it is not a restriction to consider only separating families.

Theorem 22 (Falgas-Ravry [20]) *Let $\mathcal{A}$ be a separating union-closed family on m elements. Then*

$$\frac{1}{m} \cdot \sum_{u \in U(\mathcal{A})} |\mathcal{A}_u| \geq \frac{m+1}{2}. \tag{7}$$

He remarks that this bound is stronger than Reimer's bound if $m > \sqrt{n \log_2 n}$. The proof of (7) is rather simple:

Proof Assume that the elements $1, 2, \ldots, m$ of $U(\mathcal{A})$ are labelled in order of increasing frequency. As $\mathcal{A}$ is separating, this ordering ensures that for any $1 \leq i < j \leq m$ there is a set $X_{ij} \in \mathcal{A}$ such that $i \notin X_{ij}$ and $j \in X_{ij}$. For all $1 \leq i \leq m - 1$ let $X_i = \bigcup_{j=i+1}^{m} X_{ij}$, and put $X_0 := U(\mathcal{A})$. Observe that (a) the X_i are all distinct and that (b) $[i + 1, m] \subseteq X_i$. Thus, the statement follows from

$$\sum_{u \in U(\mathcal{A})} |\mathcal{A}_u| \overset{(a)}{\geq} \sum_{i=0}^{m-1} |X_i| \overset{(b)}{\geq} \sum_{i=0}^{m-1} (m - i) = \frac{m(m+1)}{2}.$$

$\square$

Let us point out an easy consequence of the proof. As Nishimura and Takahashi observed, the union-closed sets conjecture holds for families that are very large with respect to their universe. Here we obtain the analogous result for very *small* families:

Theorem 23 *Any separating family on m elements with at most 2m member-sets satisfies the union-closed sets conjecture.*

Proof Each of the m sets X_i as constructed above contains the most frequent element x_m. $\square$

We note that this is a weaker bound than the one obtained by Lo Faro for a minimal counterexample (Lemma 18): $n \leq 4m - 1$. However, Lo Faro's techniques do not extend easily to small families and there is a good reason for this. If the factor in Theorem 23 can be improved to $c > 2$ then we may deduce that there is always an element whose frequency is a constant fraction of the number of member-sets. This natural weakening of the union-closed sets conjecture is still very much open.

Theorem 24 (Hu [30]) *Suppose there is a $c > 2$ so that any separating union-closed family $\mathcal{A}'$ with $|\mathcal{A}'| \leq c|U(\mathcal{A}')|$ satisfies the union-closed sets conjecture. Then, for every union-closed family $\mathcal{A}$, there is an element u of frequency*

$$|\mathcal{A}_u| \geq \frac{c-2}{2(c-1)}|\mathcal{A}|.$$

The theorem is proved along the following lines: by cloning some element, the universe U of $\mathcal{A}$ is enlarged to U'. At the same time, we add sets of the form $U' - x$ in order to separate the clones from each other. The resulting family $\mathcal{A}'$ is then separating and will be made to have size $|\mathcal{A}'| \leq c|U'|$. Now an element of frequency $\geq \frac{1}{2}|\mathcal{A}'|$ will still have high frequency in $\mathcal{A}$.

Falgas-Ravry also gives a family of separating union-closed families which shows that the combination of the bounds (5) and (7) is close to optimal, in the sense that the sum of both bounds can serve as an upper bound on the minimum possible weight of a separable union-closed family. For this, he calls a pair (m, n) *satisfiable* if there is a separating union-closed family with n sets on a universe of m elements.

Theorem 25 (Falgas-Ravry [20] and Reimer [51]) *Let (m, n) be a satisfiable pair of integers. Let $\mathcal{A}$ be a union-closed family on m elements and n sets of minimal average frequency. Then*

$$\max\left(\frac{n \log_2 n}{2m}, \frac{m+1}{2}\right) \leq \frac{1}{m} \cdot \sum_{u \in U(\mathcal{A})} |\mathcal{A}_u| \leq \frac{n \log_2 n}{2m} + \frac{m+1}{2} + \frac{n}{m}. \quad (8)$$

To establish the upper bound in Theorem 25, Falgas-Ravry uses a construction not unlike that of Duffus and Sands [18] that we discuss below.

6.3 Limits of Averaging

In the framework of the lattice formulation, Czédli et al. [14] construct for every size m of the universe a family of $\lfloor \frac{2}{3}2^m \rfloor$ members, for which averaging fails. We present here a lattice-free version of their family and a short and elementary proof that the average is always too small.

On the set $\mathbb{N}^{<\omega}$ of finite subsets of the positive integers, let $<$ be the order defined by first sorting by increasing largest element and then by reverse colex order. In other words, we set $A < B$ if

- $\max A < \max B$; or

- max $A = \max B$ but $\max(A \triangle B) \in A$

for finite $A, B \subseteq \mathbb{N}$.

As an illustration, here is the initial segment of the order, where we write 124 for the set $\{1, 2, 4\}$:

$$\emptyset < 1 < 12 < 2 < 123 < 23 < 13 < 3 < 1234 < 234$$
$$< 134 < 34 < 124 < 24 < 14 < 4 < 12345 < \ldots$$

For any positive integer n, define the *Hungarian family* $\mathcal{H}^{(n)}$ to be the inital segment of length n of $\mathbb{N}^{<\omega}$ under $<$. It is easy to check that $\mathcal{H}^{(n)}$ is union-closed and that its universe is $[\lceil \log_2 n \rceil]$.

Theorem 26 (Czédli, Maróti and Schmidt) *For the Hungarian family on $[m]$ of size $n = \lfloor \frac{2}{3} 2^m \rfloor$*

$$\frac{1}{m} \cdot \sum_{i \in [m]} \left| \mathcal{H}_i^{(n)} \right| < \frac{|\mathcal{H}^{(n)}|}{2}.$$

for any $m > 1$.

Proof The key to the proof are the simple and well-known identities

$$\left\lfloor \frac{2}{3} 2^m \right\rfloor = \frac{2^{m+1} - 1}{3} = 2^{m-1} + 2^{m-3} + \cdots + 4 + 1 \quad \text{if } m \text{ odd.} \tag{9}$$

$$\left\lfloor \frac{2}{3} 2^m \right\rfloor = \frac{2^{m+1} - 2}{3} = 2^{m-1} + 2^{m-3} + \cdots + 8 + 2 \quad \text{if } m \text{ even.} \tag{10}$$

Put $k = \lfloor \frac{m-1}{2} \rfloor$. Denote by I_0 the initial segment of $\mathbb{N}^{<\omega}$ of length 2^{m-1}, by I_1 the set of the next 2^{m-3} sets in the order, by I_2 the following 2^{m-5} sets and so on until we reach I_k.

Clearly, $|I_i| = 2^{m-(2i+1)}$ and $\mathcal{H}^{(n)} = I_0 \cup I_1 \cup \cdots \cup I_k$. Moreover, we can see that $I_0 = 2^{[m-1]}$ and that for $i \geq 1$, the set I_i is the set of all $X \subseteq [m]$ that contain all of $m - 1, m - 3, \ldots, m - (2i - 1)$ and of $m, m - 2i$, but none of $m - 2, m - 4, \ldots, m - (2i - 2)$.

Thus, an element $m - (2i - 1)$ appears in half of the members of $I_0 \cup \cdots \cup I_{i-1}$ and in all of the sets in $I_i \cup \cdots \cup I_k$. Its frequency is therefore

$$\left| \mathcal{H}_{m-(2i-1)}^{(n)} \right| = \tfrac{1}{2} (|I_0| + \cdots + |I_{i-1}|) + |I_i| + \cdots + |I_k|. \tag{11}$$

An element $m - 2i$ is contained in half of the sets of $I_0 \cup \cdots \cup I_{i-1}$, in all of the sets in I_i but in none of $I_{i+1} \cup \cdots \cup I_k$. Its frequency is

$$\left| \mathcal{H}_{m-2i}^{(n)} \right| = \tfrac{1}{2} (|I_0| + \cdots + |I_{i-1}|) + |I_i|. \tag{12}$$

Moreover, we observe that m lies in all of sets of $\mathcal{H}^{(n)}$ but those in I_0.

For the final argument, we assume m to be even, that is $m = 2k + 2$. The case of odd m is very similar. With (11) and (12), we obtain

$$\sum_{j=1}^{m} \left| \mathcal{H}_j^{(n)} \right| = \left| \mathcal{H}_m^{(n)} \right| + \sum_{i=1}^{k} \left(\left| \mathcal{H}_{m-(2i-1)}^{(n)} \right| + \left| \mathcal{H}_{m-2i}^{(n)} \right| \right) + \left| \mathcal{H}_1^{(n)} \right|$$

$$= |\mathcal{H}^{(n)}| - |I_0| + \sum_{i=1}^{k} \left(|\mathcal{H}^{(n)}| + |I_i| \right) + \frac{1}{2} |\mathcal{H}^{(n)}|$$

$$= (k+1)|\mathcal{H}^{(n)}| - 2|I_0| + \frac{3}{2} |\mathcal{H}^{(n)}|$$

$$= \frac{m}{2} |\mathcal{H}^{(n)}| - 2^m + \frac{3}{2} \cdot \frac{2^{m+1} - 2}{3} = \frac{m}{2} |\mathcal{H}^{(n)}| - 1,$$

where we used (10) in the penultimate step.   □

So, the averaging method can never yield the union-closed sets conjecture in its full generality. Might it perhaps be possible to at least obtain the natural relaxation, in which we only ask for an element that appears in ≥ 1 % of the member-sets? As Duffus and Sands [18] observed, not even this more modest aim may be attained just by averaging. We present here their construction.

Let V be a set of size $2t$, and $W = \{w_1, \ldots, w_{2^t}\}$ be a disjoint set of 2^t elements. Put

$$\mathcal{A} = 2^V \cup \{V \cup \{w_1, \ldots, w_i\} \colon i = 1 \ldots, 2^t\}.$$

Then $\mathcal{A}$ is a (separating) union-closed family of size $|\mathcal{A}| = 2^{2t} + 2^t$ on a universe $U = V \cup W$ of size $2t + 2^t$. Averaging yields

$$\frac{1}{|U|} \cdot \sum_{u \in U} \frac{|\mathcal{A}_u|}{|\mathcal{A}|} = \frac{2t(2^{2t-1} + 2^t) + \sum_{i=1}^{2^t}(2^t - i + 1)}{(2t + 2^t)(2^{2t} + 2^t)}$$

$$= \frac{2t(2^{2t-1} + 2^t) + 2^{t-1}(2^t - 1)}{(2t + 2^t)(2^{2t} + 2^t)} \to 0 \quad \text{as } t \to \infty,$$

as the largest summand in the numerator is $t2^{2t}$, while the largest one in the denominator is 2^{3t}. This shows that an averaging argument cannot always guarantee an element of frequency at least $c|\mathcal{A}|$ for any $c > 0$.

6.4 Up-Compression

We now outline Reimer's proof of Theorem 21 because it uses a common technique in extremal combinatorics: shifting or compression. We first restate the theorem.

Theorem 21 (Reimer [51]) *Let $\mathcal{A}$ be a union-closed family on n sets. Then*

$$\frac{1}{n} \cdot \sum_{A \in \mathcal{A}} |A| \geq \frac{\log_2 n}{2}.$$

Compression subjects the given initial object (the union-closed family), to small incremental changes until a simpler object is reached (an up-set), while maintaining the essential properties of the initial object. Variants of compression have been used by Frankl in order to prove the Kruskal–Katona theorem [25] and in the context of traces of finite sets [24]. The technique is also used by Alon [5] and various others; see Kalai's blog post [34] for an enlightening discussion.

Returning to Reimer's proof we define the *up-compression* of a union-closed family $\mathcal{A}$. For this, consider an element i, and define

$$u_i(A) = \begin{cases} A + i & \text{if } A + i \notin \mathcal{A} \\ A & \text{otherwise,} \end{cases}$$

for every $A \in \mathcal{A}$. Then it turns out that the up-compressed family $u_i(\mathcal{A}) := \{u_i(A) : A \in \mathcal{A}\}$ is still union-closed. Moreover, iteratively applying up-compression for every element i in the universe of $\mathcal{A}$ results in an *up-set*: a family $\mathcal{U}$ on universe U for which $X \in U$ and $X \subseteq Y \subseteq U$ implies $Y \in \mathcal{U}$. We may always assume $\mathcal{A}$ to have universe $[m]$. We then write $u(\mathcal{A})$ for the iterated up-compression $u_m \circ \cdots \circ u_1(\mathcal{A})$.

Lemma 27 (Reimer [51]) *Let $\mathcal{A}$ be a union-closed family on universe U. Then*

(i) $u_i(\mathcal{A})$ *is union-closed for any $i \in U$; and*
(ii) $u(\mathcal{A})$ *is an up-set.*

What have we gained? The key to the averaging technique is to control the total set size $\sum_{A \in \mathcal{A}} |A|$. For an up-set the total set size can be given in a closed form. Define the *edge boundary* of an up-set $\mathcal{U}$ on a universe U to be

$$EB(\mathcal{U}) = \{(A, A + i) \colon A \notin \mathcal{U}, \ i \in U \text{ and } A + i \in \mathcal{U}\}.$$

Now

Lemma 28 (Reimer [51]) *Let $\mathcal{U}$ be an up-set on m elements. Then*

$$2 \sum_{A \in \mathcal{U}} |A| = m|\mathcal{U}| + |EB(\mathcal{U})|.$$

In order to finish Reimer's proof we need to see that the second essential part of the compression argument holds: that the object does not change too much during compression. Here this means that the total set size has controlled growth.

Lemma 29 (Reimer [51]) *Let $\mathcal{A}$ be union-closed family. Then*

(i) $\sum_{A \in \mathcal{A}} |u(A) - A| \leq |EB(u(\mathcal{A}))|$; *and*

(ii) $\sum_{A \in \mathcal{A}} |u(A) - A| \leq |\mathcal{A}|(m - \log_2(|\mathcal{A}|))$.

Proof of Theorem 21 Applying the previous lemmas we obtain

$$
\begin{aligned}
2 \sum_{A \in \mathcal{A}} |A| &= 2 \sum_{A \in \mathcal{A}} |u(A)| - 2 \sum_{A \in \mathcal{A}} |u(A) - A| \\
&\geq m|u(\mathcal{A})| + |EB(u(\mathcal{A}))| - 2 \sum_{A \in \mathcal{A}} |u(A) - A| \\
&\geq m|\mathcal{A}| + |EB(u(\mathcal{A}))| - |EB(u(\mathcal{A}))| - |\mathcal{A}|(m - \log_2(|\mathcal{A}|)) \\
&= |\mathcal{A}| \cdot \log_2(|\mathcal{A}|).
\end{aligned}
$$

$\square$

Refining Reimer's approach, Balla, Bollobás and Eccles improve substantially on Nishimura and Takahashi's observation that large union-closed families never pose a counterexample to Frankl's conjecture.

Theorem 30 (Balla et al. [9]) *Any union-closed family on m elements with at least $\lceil \frac{2}{3} 2^m \rceil$ member-sets satisfies the union-closed sets conjecture.*

In fact, Balla et al. prove that the average frequency of such a family $\mathcal{A}$ is always at least $\frac{|\mathcal{A}|}{2}$. In view of Theorem 26 this is best possible.

The key idea of the proof of Theorem 30 is to exploit the Kruskal–Katona theorem in conjunction with up-compression. This allows to show that, among all union-closed families on n member-sets, the Hungarian family $\mathcal{H}^{(n)}$ has minimal total set size. Since the total set size of $\mathcal{H}^{(n)}$ is large, provided that $n \geq \lceil \frac{2}{3} 2^m \rceil$, the double-counting argument (4) then yields an average frequency that is large enough to imply the union-closed sets conjecture for the given family.

Up-compression, and in particular, the effect of the order in which the elements i of the universe are chosen for the up-compression is further investigated by Rodaro [59]. In a fairly involved article with a heavy algebraic flavour he arrives at an upper-bound on the number of basis sets of the union-closed family. (Recall that a non-empty $B \in \mathcal{A}$ is a basis set if $B = A \cup A'$ for $A, A' \in \mathcal{A}$ implies $A = B$ or $A' = B$.) Rodaro's bound, however, is weaker than a result of Kleitman from 1976 on set families that are union-free. Cast in the language of basis sets of a union-closed family the result becomes:

Theorem 31 (Kleitman [35]) *Let $\mathcal{A}$ be a union-closed family on m elements. Then the number of basis sets is at most*

$$
\binom{m}{\lfloor \frac{m}{2} \rfloor} + \frac{2^m}{m}.
$$

While it is not clear how sharp the bound is, a family with $\binom{m}{\lfloor \frac{m}{2} \rfloor}$ basis sets is easily found: simply take all subsets of $2^{[m]}$ of size at least $\lfloor \frac{m}{2} \rfloor$.

Up-compression is clearly a powerful concept. So, it seems enticing to apply the method in a more direct way to attack Frankl's conjecture: given a union-closed family $\mathcal{A}$, choose an element i in its universe and apply up-compression with respect to i, and then reduce the problem to the hopefully simpler family $u_i(\mathcal{A})$. Unfortunately, the up-compressed family $u_i(\mathcal{A})$ is much too simple with respect to the union-closed sets conjecture: the family satisfies it for trivial reasons. Indeed, the element i always appears in at least half of the member-sets of $u_i(\mathcal{A})$.

Faro [22] found a way to circumvent this. Call an element y *dominated by* x if $y \in A \in \mathcal{A}$ implies $x \in A$—in other words, when $\mathcal{A}_y \subseteq \mathcal{A}_x$. Then we may apply up-compression with respect to y selectively to the sets in $\mathcal{A}_x$. That is, we set

$$u'_y(A) := \begin{cases} A + y & \text{if } A \in \mathcal{A}_x \text{ and } A + y \notin \mathcal{A} \\ A & \text{otherwise.} \end{cases}$$

The resulting family $\mathcal{A}' := u'_y(\mathcal{A})$ is still union-closed. Moreover, the frequency of y is bounded by the frequency of x, which has not changed. If $\mathcal{A}'$ satisfies the union-closed sets conjecture then this is also the case for the original family $\mathcal{A}$. Thus, this restricted up-compression allows to force more structure without augmenting the frequency. While Lo Faro manages to exploit this technique in order to obtain a bound on a minimal counterexample it is not clear whether it or a variant may be used to a more far-reaching effect.

We note that up-compression is also used by Leck and Roberts [39] in the context of the union-closed sets conjecture.

6.5 Generalised Averages

We saw in the previous section that the Hungarian family $\mathcal{H}^{(n)}$ has minimum total set size among all union-closed families with n member-sets. Leck et al. [40] study a more general set-up, in which they allow the set sizes to be weighted. For this, they consider non-negative weight functions $w \colon 2^{[m]} \to \mathbb{R}_{\geq 0}$ that are constant on all sets of the same size. That is, there are reals $w_i \geq 0$ so that $w(X) = w_i$ if $|X| = i$, for every $X \subseteq [m]$. Moreover, the weights are non-decreasing with i, meaning $w_0 \leq w_1 \leq \cdots \leq w_m$. The *weight* of a non-empty union-closed family $\mathcal{A}$ is then defined as $\sum_{A \in \mathcal{A}} w(A)$. For example, if $w_i = i$ for all $i \in [0, m]$, then $w(\mathcal{A})$ is just the total set size.

For families generated by 2-sets, Leck et al. managed to determine the extremal families. These families turn out to be independent of the actual weight. In contrast to above, where we used the reverse colex order we need here the standard colex order: if $X, Y \subseteq [m]$ are distinct then $X < Y$ if and only if $\max(X \Delta Y) \in Y$. Then, we define $\mathcal{U}_k$ to be the union-closure of the first k distinct 2-sets in the colex order. For any weight w, Leck et al. calculate the weight of $\mathcal{U}_k$ to be

$$\sum_{i=2}^{a+2} \left(\binom{a+1}{i} - \binom{a-b}{i-1} \right) \cdot w_i,$$

where a and b are any integers such that $0 \leq b \leq a$ and $k = \binom{a}{2} + b$.

Theorem 32 (Leck et al. [40]) *For every k and every weight w, the family $\mathcal{U}_k$ has minimum weight $w(\mathcal{U}_k)$ among all union-closed families generated by k distinct 2-sets.*

A partial result of this had already been proved by Imrich et al. [32], first mentioned in their technical report [31], which showed that any union-closed family $\mathcal{A}$ that is generated by basis sets of size 2, has an average set size of at least $\frac{1}{2}|U(\mathcal{A})|$.

As we observed in Sect. 6.3, averaging does not always succeed, that is, the arithmetic mean of the frequencies is sometimes too low to conclude that the union-closed sets conjecture holds for a given family. For some families, such as the Hungarian family discussed above, this is because there is one or perhaps a few elements with very low frequency. Those elements might be so rare that, on the whole, the average frequency drops below the Frankl threshold of half of the member-sets.

One way to overcome this obstacle is to use a different mean than the arithmetic mean, one that de-emphasises the weight of extremely rare outliers. This approach has been pursued by Duffus and Sands [18]. While they consider a quasi-arithmetic mean for the lattice formulation, we present here the equivalent form in the set formulation. In particular, Duffus and Sands pose the question whether there is a $c > 1$ so that

$$\frac{1}{|U|} \sum_{u \in U} c^{|\mathcal{A}_u|} \geq c^{\frac{|\mathcal{A}|}{2}} \tag{13}$$

for all union-closed families $\mathcal{A}$ with universe U. Clearly, (13) would imply the union-closed sets conjecture. As evidence, Duffus and Sands prove that the lattice version of (13) holds for distributive lattices when $c = 4$.

While (13) seems quite enticing, a new idea is needed to make this, or some other, generalised average work. Indeed, it is no longer obvious how the main advantage of the averaging approach can be exploited, namely that the frequencies are analysed *indirectly* via the set sizes. In the case of distributive lattices, Duffus and Sands could investigate the individual frequencies $|\mathcal{A}_u|$ to arrive at their result. In general, this will not be possible. For, if it was, then there would be no need to consider a quasiarithmetic mean (or of any other kind), as one could immediately exhibit an abundant element.

6.6 Families of Minimum Density

Rather than averaging the frequencies over the whole universe, we may hope to gain more by restricting the range of the average, for example to the elements of the smallest member-set. This approach was developed by Wójcik [70] and followed up by Balla [8].

Define s_k to be the largest real so that for any union-closed family $\mathcal{A}$ and any k-set S in $\mathcal{A}$ it holds that

$$\frac{1}{|S|} \sum_{u \in S} |\mathcal{A}_u| \geq s_k |\mathcal{A}|. \tag{14}$$

The first 10 values have been determined exactly by Wójcik; we list here the first five: $s_1 = \frac{1}{2}$, $s_2 = \frac{1}{2}$, $s_3 = \frac{4}{9}$, $s_4 = \frac{2}{5}$ and $s_5 = \frac{9}{25}$. So, in particular, any 5-set in any union-closed family will always contain an element that appears in at least a third of the member-sets.

Somewhat surprisingly, the value s_k coincides with the so-called minimal *density* of a family on k elements:

Theorem 33 (Wójcik [70]) *For every* $k \in \mathbb{N}$ *it holds that*

$$s_k = \min_{\mathcal{A}} \frac{1}{k|\mathcal{A}|} \cdot \sum_{u \in U(\mathcal{A})} |\mathcal{A}_u|,$$

where the minimum ranges over all union-closed families $\mathcal{A}$ *with* $|U(\mathcal{A})| = k$.

We mention that we have reversed here definition and consequence, as Wójcik defines the s_k as minimal densities but then proves the equivalence to (14).

Wójcik conjectured and Balla proved that:

Theorem 34 (Balla [8]) *For all* k, $s_k \geq \frac{\log_2 k}{2k}$.

The main step in the proof is an application of Reimer's theorem. As Wójcik [70] indicated, this lower bound is asymptotically optimal. To see this, consider the family $2^{[r]} \cup [k]$, where $r = \lceil \log_2 k \rceil$, and observe that its density is $(1 + o(1))\frac{\log_2 k}{2k}$. Note, however, that this family is not separating.

Combining Theorems 33 and 34, Balla arrives at a lower bound on the maximum frequency in terms of the size of the universe.

Corollary 35 (Balla [8]) *In every union-closed family on* $m \geq 16$ *elements and* n *sets there is an element contained in at least* $\sqrt{\frac{\log_2 m}{m}} \cdot \frac{n}{2}$ *many member-sets.*

7 Further Results

Sarvate and Renaud [62] observed that if the union-closed sets conjecture holds for union-closed families on n sets, n odd, then it holds for union-closed families with $n+1$ sets. In particular, n_0 is odd. Faro [22] and later Roberts and Simpson [58] proved $n_0 \geq 4m_0 - 1$. As discussed earlier, this result turns out to be very useful for families on few sets.

Another result in this direction is given by Norton and Sarvate [47]: any counterexample with n_0 sets contains at least three distinct elements of frequency exactly $\frac{n_0 - 1}{2}$. Other necessary properties of counterexamples were given by Faro [21,22] and Dohmen [17].

Peng et al. [49] study what they call the *half-life* of set families. Given a set family $\mathcal{B}$ that is not necessarily union-closed, they consider the family $\bigcup^k \mathcal{B}$ defined as the family of unions of at most k sets of $\mathcal{B}$. The half-life of $\mathcal{B}$ is then the least k such that $\bigcup^k \mathcal{B}$ satisfies the assertion of the union-closed sets conjecture.

8 Extremal Frequency

Any induction proof of the union-closed sets conjecture will likely necessitate a strengthened induction hypothesis coupled with structural insight on those families with low maximum frequencies. Let us therefore look at the minimal maximum element frequency a family on a given number of sets may have.

For a union-closed family $\mathcal{A}$ define $\phi(\mathcal{A})$ to be the maximum frequency of an element of the universe, that is,

$$\phi(\mathcal{A}) = \max_{u \in U(\mathcal{A})} |\mathcal{A}_u|.$$

Let $\phi(n)$ be the minimum over all $\phi(\mathcal{A})$, where $\mathcal{A}$ is a union-closed family of $n \geq 2$ member-sets. Clearly, this allows the trivial reformulation of the union-closed sets conjecture as:

Conjecture 36 $\phi(n) \geq \frac{n}{2}$ *for all integers $n \geq 2$.*

In this way, the union-closed sets conjecture becomes a problem about an integer sequence. What can be said about this sequence $\phi(n)$? For instance, that it is a slowly growing sequence:

Lemma 37 (Renaud [53]) $\phi(n-1) \leq \phi(n) \leq \phi(n-1) + 1$ *for all $n \geq 2$.*

Renaud[3] used the lemma to compute the first 17 values of $\phi(n)$. We put $\phi(1) = 1$ so that the sequence starts from $n = 1$ on:

$$1, 1, 2, 2, 3, 4, 4, 4, 5, 6, 7, 7, 8, 8, 8, 8, 9, 10, \ldots . \tag{15}$$

Moreover, if the union-closed sets conjecture is true, then $\phi(n) = \frac{n}{2}$ if n is a power of two, and $\phi(n) > \frac{n}{2}$ otherwise, provided Poonen's conjecture (Conjecture 14) is valid as well.

Now, there is a well-known slowly growing integer sequence that coincides with $\phi(n)$ on the initial segment (15) and that, in addition, has $a(n) = \frac{n}{2}$ if and only if n is power of two. This is Conway's challenge sequence, defined by $a(1) = a(2) = 1$ and the recurrence relation

$$a(n) = a(a(n-1)) + a(n - a(n-1)).$$

See, for instance, Kubo and Vakil [38] for background on the sequence.

As Mallows [42] proved that $a(n) \geq \frac{n}{2}$ for all $n \geq 1$, it seems tempting to seek a deeper relation between $\phi(n)$ and $a(n)$, and in some sense there is one. Renaud and Fitina construct, for every n, a union-closed family whose maximum element frequency is exactly equal to $a(n)$. We discuss this construction next.

Let us define an order $<$ on $\mathbb{N}^{(<\omega)}$, the set of finite subsets of $\mathbb{N}$, by first sorting by largest element, then by decreasing cardinality and finally by colex order. Thus, $A < B$ if

---

[3] We point out here that our sequence $\phi(n)$ equals Renaud's [53] $\phi(n-1)$.

- max $A <$ max B; or
- max $A =$ max B but $|A| > |B|$; or
- max $A =$ max B and $|A| = |B|$ but $\max(A \triangle B) \in B$

Omitting parentheses and commas this yields

$$\emptyset < 1 < 12 < 2 < 123 < 13 < 23 < 3 < 1234 < 124$$
$$< 134 < 234 < 14 < 24 < 34 < 4 < 12345 < \dots$$

as initial segment. It is easy to see that $A \leq C$ and $B \leq C$ implies $A \cup B \leq C$, which means that the first n sets of this order form a union-closed family, denoted by $\mathcal{R}(n)$.

Theorem 38 (Renaud and Fitina [55]) *For every $n \geq 2$, the most frequent element of the Renaud-Fitina family $\mathcal{R}(n)$ has frequency $a(n)$, that is,*

$$\phi(n) \leq a(n).$$

So, is $\phi(n)$ always equal to $a(n)$? By Mallows' result, that would clearly prove the union-closed sets conjecture. Unfortunately, this is not the case. In a subsequent paper, Renaud [54] described families $\mathcal{B}(n)$ whose element frequency is sometimes strictly smaller than Conways' challenge sequence. This happens for the first time at $n = 23$, where $a(n) = 14$. However, no element in the family

$$\mathcal{B}(23) = 2^{[4]} \cup \{12345, 1235, 1245, 1345, 2345, 125, 345\}$$

on 23 member-sets appears more often than 13 times. We omit the precise construction of $\mathcal{B}(n)$ but mention that it only differs from $\mathcal{R}(n)$ in the last step, when we delete sets of the same size of the power set $2^{[m]}$. There the sets to delete are chosen in a more balanced way, so that the frequency of the elements $1, \dots, m-1$ differs by at most one.

Renaud determines the maximum frequency as follows. Let

$$n = 2^m - \sum_{i=0}^{r-1} \binom{m-1}{i} - v,$$

where $0 \leq r < m - 1$ and $0 \leq v < \binom{m-1}{r}$. Then

$$\phi(\mathcal{B}(n)) = 2^{m-1} - \sum_{i=0}^{r-2} \binom{m-1}{i} - \left\lfloor \frac{rv}{m-1} \right\rfloor.$$

Furthermore, he shows that always $\phi(\mathcal{B}(n)) \leq a(n)$. Are the families $\mathcal{B}(n)$ now truly extremal, that is $\phi(n) = \phi(\mathcal{B}(n))$ for all n? Again, this is not the case. Renaud gives the example of the family

$$\mathcal{C} = 2^{[6]} \backslash \{6, 5, 16, 25, 36, 45, 136, 245\},$$

in which the most frequent element appears in 30 member-sets. However, in $\mathcal{B}(56)$ there is an element of frequency 31.

To conclude, we do not know much, in general, about the structure of an extremal family, nor are there any convincing candidates. The only exception are power sets $\mathcal{P}$, for which holds $\phi(\mathcal{P}) = \phi(|\mathcal{P}|)$, provided the union-closed sets conjecture is true. Nevertheless, the examples in this section seem to indicate that an extremal family would have relatively few elements compared to the number of member-sets: let us call a family on n member-sets and a universe of size m *compact* if $2^{m-1} < n \leq 2^m$. For example, power sets, the Renaud-Fitina families as well as the Hungarian families are compact.

Question 39 *Is it true that for a union-closed family $\mathcal{A}$ it follows from $\phi(\mathcal{A}) = \phi(|\mathcal{A}|)$ that $\mathcal{A}$ is compact?*

An affirmative answer would be a major step towards the union-closed sets conjecture. Indeed, Reimer's bound (6) in conjunction with Theorem 17 gives:

Observation 40 *Any compact union-closed family $\mathcal{A}$ contains an element that is contained in at least $\frac{6}{13}|\mathcal{A}|$ member-sets.*

While we have arrived at the end of this survey, the union-closed sets conjecture still has a bit of a journey ahead of it. We hope it will be an exciting trip.

References

1. Abe, T.: Strong semimodular lattices and Frankl's conjecture. Algebra Univers. **44**, 379–382 (2000)
2. Abe, T.: Excess of a lattice. Graphs Comb. **18**, 395–402 (2002)
3. Abe, T., Nakano, B.: Frankl's conjecture is true for modular lattices. Graphs Comb. **14**, 305–311 (1998)
4. Abe, T., Nakano, B.: Lower semimodular types of lattices: Frankl's conjecture holds for lower quasi-semimodular lattices. Graphs Comb. **16**, 1–16 (2000)
5. Alon, N.: On the density of sets of vectors. Discret. Math. **46**, 199–202 (1983)
6. A much-travelled conjecture. Aust. Math. Soc. Gaz. **14/3**, 63 (1987)
7. Union-closed sets conjecture. Aust. Math. Soc. Gaz. **14/4**, 99 (1987)
8. Balla, I.: Minimum densities of union-closed families (2011). arXiv:1106.0369v1 [math.CO]
9. Balla, I., Bollobas, B., Eccles, T.: Union-closed families of sets. J. Comb. Theory (Ser. A) **120**, 531–544 (2013)
10. Bošnjak, I., Marković, P.: The 11-element case of Frankl's conjecture. Eur. J. Comb. **15**, R88 (2008)
11. Bruhn, H., Charbit, P., Schaudt, O., Telle, J.A.: The graph formulation of the union-closed sets conjecture (2013). (Preprint)
12. Bruhn, H., Schaudt, O.: The union-closed sets conjecture almost holds for almost all random bipartite graphs (2012). (Preprint)
13. Czédli, G.: On averaging Frankl's conjecture for large union-closed sets. J. Comb. Theory (Ser. A) **116**, 724–729 (2009)
14. Czédli, G., Maróti, M., Schmidt, E.T.: On the scope of averaging for Frankl's conjecture. Order **26**, 31–48 (2009)

15. Czédli, G., Schmidt, E.T.: Frankl's conjecture for large semimodular and planar semimodular lattices. Acta Univ. Palacki. Olomuc. Fac. Rer. Nat. Math. **47**, 47–53 (2008)
16. Diestel, R.: Graph theory, 3rd edn. Springer, New York (2005)
17. Dohmen, K.: A new perspective on the union-closed sets conjecture. Ars Comb. **58**, 183–185 (2001)
18. Duffus, D., Sands, B.: An inequality for the sizes of prime filters of finite distributive lattices. Discret. Math. **201**, 89–99 (1999)
19. El-Zahar, M.: A graph-theoretic version of the union-closed sets conjecture. J. Graph Theory **26**, 155–163 (1997)
20. Falgas-Ravry, V.: Minimal weight in union-closed families. Electron. J. Comb. **18**, #P95 (2011)
21. Faro, G.L.: A note on the union-closed sets conjecture. J. Aust. Math. Soc. (Ser. A) **57**, 230–236 (1994)
22. Faro, G.L.: Union-closed sets conjecture: improved bounds. J. Comb. Math. Comb. Comput. **16**, 97–102 (1994)
23. Frankl, P.: Personal communication
24. Frankl, P.: On the trace of finite sets. J. Comb. Theory (Ser. A) **34**, 41–45 (1983)
25. Frankl, P.: A new short proof for the Kruskal–Katona theorem. Discret. Math. **48**, 327–329 (1984)
26. Frankl, P.: Handbook of combinatorics, vol. 2, pp. 1293–1329. MIT Press, Cambridge (1995)
27. Gao, W., Yu, H.: Note on the union-closed sets conjecture. Ars Comb. **49**, 280–288 (1998)
28. Grätzer, G.: General lattice theory. Springer, New York (2003)
29. Herrmann, C., Langsdorf, R.: Frankl's conjecture for lower semimodular lattices (1999). (Unpublished preprint)
30. Hu, Y.: Master's thesis (in preparation)
31. Imrich, W., Sauer, N., Woess, W.: The average size of admissible sets in a graph. Tech. report, Montanuniversität Leoben (1988)
32. Imrich, W., Sauer, N., Woess, W.: The average size of nonsingular sets in a graph. In: Sauer, N.W. (eds.) Finite and Infinite Combinatorics in Sets and Logic, pp. 199–205. Kluwer Academic Publishers, Dordrecht (1993)
33. Johnson, R.T., Vaughan, T.P.: On union-closed families, I. J. Comb. Theory (Ser. A) **85**, 112–119 (1999)
34. Kalai, G.: Extremal combinatorics IV: shifting. http://gilkalai.wordpress.com/2008/10/06/extremal-combinatorics-iv-shifting/. Accessed 01 May 2013
35. Kleitman, D.J.: Extremal properties of collections of subsets containing no two sets and their union. J. Comb. Theory (Ser. A) **20**, 390–392 (1976)
36. Knill, E.: Generalized degrees and densities for families of sets. PhD thesis, University of Colorado (1991)
37. Knill, E.: Graph generated union-closed families of sets (1994). arXiv:math/9409215v1 [math.CO]
38. Kubo, T., Vakil, R.: On Conway's recursive sequence. Discret. Math. **152**, 225–252 (1996)
39. Leck, U., Roberts, I.T.: Inequalities for cross-unions of collections of finite sets (2013). (Preprint)
40. Leck, U., Roberts, I.T., Simpson, J.: Minimizing the weight of the union-closure of families of two-sets. Australas. J. Comb. **52**, 67–73 (2012)
41. Llano, B., Montellano-Ballesteros, J.J., Rivera-Campo, E., Strausz, R.: On conjectures of Frankl and El-Zahar. J. Graph Theory **57**, 344–352 (2008)
42. Mallows, C.I.: Conway's challenge sequence. Am. Math. Mon. **98**, 5–20 (1991)
43. Marić, F., Živković, M., Vučković, B.: Formalizing Frankl's conjecture: Fc-families. Lect. Notes Comput. Sci. **7362**, 248–263 (2012)
44. Marković, P.: An attempt at Frankl's conjecture. Publ. de l'Inst. Math. Nouv. Sér. **81**, 29–43 (2007)
45. Morris, R.: FC-families, and improved bounds for Frankl's conjecture. Eur. J. Comb. **27**, 269–282 (2006)
46. Nishimura, T., Takahashi, S.: Around Frankl conjecture. Sci. Rep. Yokohama Nat. Univ. Sect. I Math. Phys. Chem. **43**, 15–23 (1996)
47. Norton, R.M., Sarvate, D.G.: A note on the union-closed sets conjecture. J. Aust. Math. Soc. (Ser. A) **55**, 411–413 (1993)
48. Open Problem Garden: Frankl's union-closed sets conjecture. http://www.openproblemgarden.org/op/frankls_union_closed_sets_conjecture. Accessed 06 May 2013
49. Peng, Y., Sissokho, P., Zhao, C.: An extremal problem for set families generated with the union and symmetric difference operations. J. Comb. **3**, 651–668 (2012)
50. Poonen, B.: Union-closed families. J. Comb. Theory (Ser. A) **59**, 253–268 (1992)
51. Reimer, D.: An average set size theorem. Comb. Probab. Comput. **12**, 89–93 (2003)

52. Reinhold, J.: Frankl's conjecture is true for lower semimodular lattices. Graphs Comb. **16**(1), 115–116 (2000)
53. Renaud, J.-C.: Is the union-closed sets conjecture the best possible? J. Aust. Math. Soc. (Ser. A) **51**, 276–283 (1991)
54. Renaud, J.-C.: A second approximation to the boundary function on union-closed collections. Ars Comb. **41**, 177–188 (1995)
55. Renaud, J.-C., Fitina, L.F.: On union-closed sets and Conway's sequence. Bull. Aust. Math. Soc. **47**, 321–332 (1993)
56. Rival, I. (ed.): Graphs and order. NATO ASI Series, vol. 147. Springer, Netherlands (1985)
57. Roberts, I.: The union-closed sets conjecture. Tech. Report 2/92, Curtin University of Technology (1992)
58. Roberts, I., Simpson, J.: A note on the union-closed sets conjecture. Australas. J. Comb. **47**, 265–267 (2010)
59. Rodaro, E.: Union-closed vs upward-closed families of finite sets (2012). arXiv:1208.5371v2 [math.CO]
60. Salvi, N.Z.: An equivalent formulation of the union-closed sets conjecture. (Manuscript)
61. Salzborn, F.: A note on the intersecting sets conjecture (1989). (Manuscript)
62. Sarvate, D.G., Renaud, J.-C.: On the union-closed sets conjecture. Ars Comb. **27**, 149–154 (1989)
63. Sarvate, D.G., Renaud, J.-C.: Improved bounds for the union-closed sets conjecture. Ars Comb. **29**, 181–185 (1990)
64. Stanley, R.P.: Enumerative combinatorics, vol. I. Wadsworth & Brooks/Cole, Pacific Grove (1986)
65. Vaughan, T.P.: More on 3-sets in union-closed families: the end is in sight. http://atlas-conferences.com/c/a/q/a/23.htm. Accessed: 24 March 2013
66. Vaughan, T.P.: Families implying the Frankl conjecture. Eur. J. Comb. **23**, 851–860 (2002)
67. Vaughan, T.P.: Three-sets in a union-closed family. J. Comb. Math. Comb. Comput. **49**, 73–84 (2004)
68. West, D.: Union-closed sets conjecture (1979). http://www.math.uiuc.edu/west/openp/unionclos.html. Accessed 06 May 2013
69. Wikipedia: Union-closed sets conjecture. http://en.wikipedia.org/wiki/Union-closed_sets_conjecture. Accessed 06 May 2013
70. Wójcik, P.: Density of union-closed families. Discret. Math. **105**, 259–267 (1992)
71. Wójcik, P.: Union-closed families of sets. Discret. Math. **199**, 173–182 (1999)
72. Živković, M., Vučković, B.: The 12 element case of Frankl's conjecture (2012). (Preprint)