Union-closed families with small average overlap densities

David Ellis

School of Mathematics
University of Bristol
Bristol, U.K.

david.ellis@bristol.ac.uk

Submitted: Dec 24, 2020; Accepted: Oct 17, 2021; Published: Jan 28, 2022

Abstract

In this very short paper, we show that the average overlap density of a union-closed family $\mathcal{F}$ of subsets of $\{1, 2, \ldots, n\}$ may be as small as

$$\Theta((\log_2 \log_2 |\mathcal{F}|)/(\log_2 |\mathcal{F}|)),$$

for infinitely many positive integers n.

Mathematics Subject Classifications: 05D05

1 Introduction

If X is a set, a family $\mathcal{F}$ of subsets of X is said to be *union-closed* if the union of any two sets in $\mathcal{F}$ is also in $\mathcal{F}$. The celebrated Union-Closed Conjecture (a conjecture of Frankl [2]) states that if X is a finite set and $\mathcal{F}$ is a union-closed family of subsets of X (with $\mathcal{F} \neq \{\varnothing\}$), then there exists an element $x \in X$ such that x is contained in at least half of the sets in $\mathcal{F}$. Despite the efforts of many researchers over the last forty-five years, and a recent Polymath project [5] aimed at resolving it, this conjecture remains wide open. It has only been proved under very strong constraints on the ground-set X or the family $\mathcal{F}$; for example, Balla, Bollobás and Eccles [1] proved it in the case where $|\mathcal{F}| \geqslant \frac{2}{3}2^{|X|}$; more recently, Karpas [3] proved it in the case where $|\mathcal{F}| \geqslant (\frac{1}{2} - c)2^{|X|}$ for a small absolute constant $c > 0$; and it is also known to hold whenever $|X| \leqslant 12$ or $|\mathcal{F}| \leqslant 50$, from work of Vučković and Živković [8] and of Roberts and Simpson [7].

In 2016, a Polymath project [5] was convened to tackle the Union-Closed Conjecture. While it did not result in a proof of the conjecture, several interesting related conjectures were posed. Among them was the 'average overlap density conjecture'.

If X is a finite set and $\mathcal{F} \subset \mathcal{P}(X)$ with $\mathcal{F} \neq \varnothing$, we define the *abundance* of x (with respect to $\mathcal{F}$) by $\gamma_x = |\{A \in \mathcal{F} : \ x \in A\}|/|\mathcal{F}|$, i.e., γ_x is the probability that a uniformly random member of $\mathcal{F}$ contains x. A natural first quantity to consider, in trying to prove the Union-Closed Conjecture, is the average abundance of a uniformly random element of the ground set, i.e., $\mathbb{E}_{x \in X}[\gamma_x]$; if this quantity were always at least $1/2$, the Union-Closed Conjecture would immediately follow. A moment's thought shows that this is false, however, e.g. by considering the union-closed family $\{\varnothing, \{1\}, \{1,2,3\}\} \subset \mathcal{P}(\{1,2,3\})$, which has average abundance $4/9$. Similarly, for any $n \in \mathbb{N}$, the union-closed family $\mathcal{F} = \{\varnothing, \{1\}, \{1,2\}, \ldots, \{1,2,\ldots,\lfloor\sqrt{n}\rfloor\}, \{1,2,3,\ldots,n\}\} \subset \mathcal{P}(\{1,2,\ldots,n\})$ has average abundance $\Theta(1/\sqrt{n}) = \Theta(1/|\mathcal{F}|)$.

It is natural to consider the expected abundance of a random element of the ground-set X chosen according to other (non-uniform) distributions on X. The following was considered in the Polymath project [5]. We define the *average overlap density* $\mathrm{AOD}(\mathcal{F})$ of $\mathcal{F}$ to be the expected value of γ_x, where x is a uniformly random element of a uniformly random nonempty member of $\mathcal{F}$:

$$\begin{aligned}
\mathrm{AOD}(\mathcal{F}) &:= \frac{1}{|\mathcal{F} \setminus \{\varnothing\}|} \sum_{A \in \mathcal{F} \setminus \{\varnothing\}} \frac{1}{|A|} \sum_{x \in A} \gamma_x \\
&= \frac{1}{|\mathcal{F} \setminus \{\varnothing\}|} \sum_{A \in \mathcal{F} \setminus \{\varnothing\}} \frac{1}{|A|} \sum_{x \in A} \frac{|\{B \in \mathcal{F} : \ x \in B\}|}{|\mathcal{F}|} \\
&= \frac{1}{|\mathcal{F} \setminus \{\varnothing\}|} \sum_{A \in \mathcal{F} \setminus \{\varnothing\}} \left(\frac{1}{|\mathcal{F}|} \sum_{B \in \mathcal{F}} \frac{|A \cap B|}{|A|} \right) \\
&= \mathbb{E}_{A \in \mathcal{F} \setminus \{\varnothing\}} \mathbb{E}_{B \in \mathcal{F}} \left[\frac{|A \cap B|}{|A|} \right]. \qquad (1)
\end{aligned}$$

(The first and second expectations in (1) are of course over a uniformly random element of $\mathcal{F} \setminus \{\varnothing\}$, and a uniformly random element of $\mathcal{F}$, respectively.) The last equality justifies the 'average overlap' terminology. The average overlap density conjecture stated that if X is a finite set, and $\mathcal{F}$ is a union-closed family of subsets of X with $\mathcal{F} \neq \varnothing$ and $\mathcal{F} \neq \{\varnothing\}$, then the average overlap density of $\mathcal{F}$ is at least $1/2$. Clearly, it would immediately imply the Union-Closed Conjecture.

Unfortunately, the average overlap density conjecture was quickly shown to be false (during the Polymath project [6]); an infinite sequence of union-closed families $\mathcal{F}_n \subset \mathcal{P}(\{1,2,\ldots,n\})$ was constructed with $\mathrm{AOD}(\mathcal{F}_n) = 7/15 + o(1)$ as $n \to \infty$. However, the following weakening of the average overlap density conjecture remained open.

Conjecture 1. There exists an absolute positive constant $c > 0$ such that the following holds. Let $n \in \mathbb{N}$ and let $\mathcal{F} \subset \mathcal{P}(\{1,2,\ldots.n\})$ be union-closed with $\mathcal{F} \neq \{\varnothing\}$. Then the average overlap density of $\mathcal{F}$ is at least c.

Conjecture 1 would immediately imply the weakening of the Union-Closed Conjecture where $1/2$ is replaced by the absolute positive constant c. In this very short paper, we prove the following.

Theorem 2. *For infinitely many positive integers n, there exists a union-closed family $\mathcal{F}$ of subsets of $\{1,2,\ldots,n\}$ whose average overlap density is $\Theta((\log_2\log_2|\mathcal{F}|)/(\log_2|\mathcal{F}|))$.*

This disproves Conjecture 1 in a strong sense. It follows from an old result of Knill [4] that if $\mathcal{F} \subset \mathcal{P}(\{1,2,\ldots,n\})$ is union-closed, then there exists $x \in \{1,2,\ldots,n\}$ with abundance $\gamma_x = \Omega(1/(\log_2|\mathcal{F}|))$, so the average overlap density can, in the best-case scenario, only be used to improve this lower bound by a factor of $\Theta(\log_2\log_2|\mathcal{F}|)$.

2 Proof of Theorem 2

For $n \in \mathbb{N}$, we write $[n] := \{1,2,\ldots,n\}$ for the standard n-element set, and if $\mathcal{G} \subset \mathcal{P}(X)$, the *union-closed family generated by* $\mathcal{G}$ is defined to be the smallest union-closed family of subsets of X that contains $\mathcal{G}$.

Let $k,m,s \in \mathbb{N}$ with $s \leqslant k-2$ and $m \geqslant 2$, and let $n = km$. Partition $[n]$ into m sets $B_1,\ldots,B_m$ with $|B_i| = k$ for all i; in what follows, we will refer to the B_i as 'blocks'. For each $i \in [m]$, choose a subset $T_i \subset B_i$ with $|T_i| = s$, and let $T = \cup_{i=1}^m T_i$. Now let $\mathcal{F} \subset \mathcal{P}([n])$ be the union-closed family generated by $\{B_i \cup \{j\} : \ i \in [m],\ j \in T\}$. Note that every set in $\mathcal{F}$ contains at least one block. The number of sets in $\mathcal{F}$ containing exactly one block is $m2^{(m-1)s}$, and in general, for each $j \in [m]$, number N_j of sets in $\mathcal{F}$ containing exactly j blocks is $\binom{m}{j}2^{(m-j)s}$, so

$$N := |\mathcal{F}| = \sum_{j=1}^m N_j = 2^{(m-1)s}\sum_{j=1}^m \binom{m}{j}2^{-(j-1)s}.$$

For each $j \in [m]$, define $p_j := N_j/N$; this is of course the probability that a uniformly random member of $\mathcal{F}$ contains exactly j blocks. We note that

$$\frac{p_{j+1}}{p_j} = \frac{N_{j+1}}{N_j} = \frac{m-j}{j+1}2^{-s} \leqslant m2^{-s} \quad \forall j \in [m-1].$$

Write $\tau := m2^{-s}$. For any $x \in [n] \setminus T$, we clearly have

$$\gamma_x = \frac{1}{m}\sum_{j=1}^m jp_j,$$

since the conditional probability that x is contained in a random member A of $\mathcal{F}$, given that A contains exactly j blocks, is j/m. We have $p_j \leqslant \tau^{j-1}p_1$ for all $j \in [m]$, and therefore for any $x \in [n] \setminus T$, we have

$$\frac{1}{m} \leqslant \gamma_x \leqslant \frac{1}{m}(1+2\tau+3\tau^2+\ldots+m\tau^{m-1}) \leqslant \frac{1}{m}(1+4\tau) \leqslant \frac{2}{m},$$

provided $\tau = m2^{-s} \leqslant 1/4$. Now, every member A of $\mathcal{F}$ contains at least one block, so for any member A of $\mathcal{F}$, the probability a uniformly random element of A is in T, is at most

$\frac{ms}{k}$. Crudely, we have $1/2 \leqslant \gamma_x \leqslant 1$ for all $x \in T$, since $A \mapsto A \cup \{x\}$ is an injection from $\{A \in \mathcal{F} : x \notin A\}$ to $\{A \in \mathcal{F} : x \in A\}$, for any $x \in T$. Hence, we have

$$\frac{1}{m} \leqslant \mathrm{AOD}(\mathcal{F}) \leqslant \Big(1 - \frac{ms}{k}\Big) \cdot \frac{2}{m} + \frac{ms}{k} \cdot 1 \leqslant \frac{2}{m} + \frac{m^2 s}{n}, \tag{2}$$

again provided $\tau = m2^{-s} \leqslant 1/4$. Now we wish to minimize the right-hand side of (2), subject to the constraint $m2^{-s} \leqslant 1/4$; clearly the optimal choice is to take $s = \lceil \log_2 m \rceil + 2$, which yields

$$\frac{1}{m} \leqslant \mathrm{AOD}(\mathcal{F}) \leqslant \frac{2}{m} + \frac{m^2 \log_2 m}{n} + O(m^2/n). \tag{3}$$

It is clear that the optimal choice of m to minimize the right-hand side of (3) is

$$m = \Theta\left(\left(\frac{n}{\log_2 n}\right)^{1/3}\right),$$

yielding $\mathrm{AOD}(\mathcal{F}) = \Theta(((\log_2 n)/n)^{1/3})$. Since, with these choices, we have

$$\log_2 |\mathcal{F}| = \Theta(n^{1/3}(\log_2 n)^{2/3}),$$

it follows that

$$\mathrm{AOD}(\mathcal{F}) = \Theta\left(\frac{\log_2 \log_2 |\mathcal{F}|}{\log_2 |\mathcal{F}|}\right),$$

proving Theorem 2.

We proceed to note two further properties of the above construction. Firstly, the average abundance of a uniformly random element of $[n]$ (with respect to $\mathcal{F}$) satisfies

$$\mathbb{E}_{x \in [n]}[\gamma_x] = \Theta\left(\frac{\log_2 \log_2 |\mathcal{F}|}{\log_2 |\mathcal{F}|}\right).$$

Secondly, the family $\mathcal{F}$ constructed above does not separate the points of $[n]$. (We say a family $\mathcal{F} \subset \mathcal{P}([n])$ *separates the points of* $[n]$ if for any $i \neq j \in [n]$ there exists $A \in \mathcal{F}$ such that $|A \cap \{i,j\}| = 1$. It is easy to see that, in attempting to prove the Union-Closed Conjecture, we may assume that the union-closed family in question separates the points of the ground set, and this assumption was adopted for much of the Polymath project [5].) However, it is easy to see that the union-closed family $\mathcal{F} \cup \{[n] \setminus \{j\} : j \in [n]\}$ has asymptotically the same average overlap density as $\mathcal{F}$ (and asymptotically the same average abundance as $\mathcal{F}$), and does separate the points of $[n]$.

References

[1] I. Balla, B. Bollobás and T. Eccles, Union-closed families of sets. *J. Combin. Theory (Series A)*, 120 (2013), 531–544.

[2] D. Duffus, in: I. Rival (Ed.), *Graphs and Order*. Reidel, Dordrecht, Boston, 1985, p. 525.

[3] I. Karpas, Two Results on Union-Closed Families. Preprint, August 2017. arXiv:1708.01434.

[4] E. Knill, Graph generated union-closed families of sets. Manuscript, September 1994. arXiv:9409215.

[5] *Polymath11: Frankl's Union-Closed Conjecture.* 1: Strengthenings, variants, potential counterexamples. https://gowers.wordpress.com/2016/01/29/func1-strengthenings-variants-potential-counterexamples/.

[6] *Polymath11: Frankl's Union-Closed Conjecture.* 2: More examples. https://gowers.wordpress.com/2016/02/08/func2-more-examples/.

[7] I. Roberts and J. Simpson, A note on the union-closed sets conjecture. *Australas. J. Combin.*, 47 (2010), 265–267.

[8] B. Vučković and M. Živković, The 12-element case of Frankl's conjecture. *IPSI Transactions on Advanced Research*, January 2017, Paper 9.