Probably Intersecting Families are Not Nested

PAUL A. RUSSELL[1] and MARK WALTERS[2]

[1]Department of Pure Mathematics and Mathematical Statistics, Centre for Mathematical Sciences,
Wilberforce Road, Cambridge CB3 0WB, UK
(e-mail: P.A.Russell@dpmms.cam.ac.uk)
[2]School of Mathematical Sciences, Queen Mary, University of London, London E1 4NS, UK
(e-mail: m.walters@qmul.ac.uk)

It is well known that an intersecting family of subsets of an n-element set can contain at most 2^{n-1} sets. It is natural to wonder how 'close' to intersecting a family of size greater than 2^{n-1} can be. Katona, Katona and Katona introduced the idea of a 'most probably intersecting family'. Suppose that $\mathcal{A}$ is a family and that $0 < p < 1$. Let $\mathcal{A}(p)$ be the (random) family formed by selecting each set in $\mathcal{A}$ independently with probability p. A family $\mathcal{A}$ is *most probably intersecting* if it maximizes the probability that $\mathcal{A}(p)$ is intersecting over all families of size $|\mathcal{A}|$.

Katona, Katona and Katona conjectured that there is a nested sequence consisting of most probably intersecting families of every possible size. We show that this conjecture is false for every value of p provided that n is sufficiently large.

AMS 2010 *Mathematics subject classification*: Primary 05D05

1. Introduction

We start by recalling the definition of an *intersecting family*: we say that $\mathcal{A} \subset \mathcal{P}([n])$ is *intersecting* if for any $A, A' \in \mathcal{A}$ we have $A \cap A' \neq \emptyset$. Since no intersecting family can contain both a set A and its complement A^c, it is easy to see that there is no intersecting family containing more than 2^{n-1} sets. We remark that this upper bound is tight and that, in fact, any intersecting family can be extended to an intersecting family of this size.

Having observed this bound, it is natural to wonder how 'close' to intersecting a family of size greater than 2^{n-1} can be. Katona, Katona and Katona [4] introduced the idea of a *most probably intersecting family*. Suppose that $\mathcal{A}$ is a family and that $0 < p < 1$. Let $\mathcal{A}(p)$ be the (random) family formed by selecting each set in $\mathcal{A}$ independently with probability p. They asked which family $\mathcal{A}$ of given size maximizes the probability that $\mathcal{A}(p)$ is intersecting.

In the same paper, they solve this problem in cases where $|\mathcal{A}|$ is only a little greater than 2^{n-1}. More precisely, they find extremal families for $|\mathcal{A}| \leqslant 2^{n-1} + \binom{n-1}{\lfloor (n-3)/2 \rfloor}$.

They also conjectured that there is a sequence of extremal families $\mathcal{A}_i$ with $|\mathcal{A}_i| = 2^{n-1} + i$ that is nested: that is, $\mathcal{A}_i \subset \mathcal{A}_j$ whenever $i < j$. In their paper it is a little unclear for which p they

make this conjecture (since there seems to be no reason to believe the optimal families are the same for different p, *a priori* it could be the case that the sequence of optimal families is nested for some values of p but not for other values). We remark that there is a simple counter-example if $p \ll 2^{-n}$ (see the discussion following Theorem 1.2 below) and so clearly this was not what was meant. In this paper we prove the following theorem, which shows that the conjecture is false for all p.

Theorem 1.1. *Suppose that $n \geqslant 21$. Let $2 \leqslant s \leqslant 2^{n-1}$. Then:*

- $[n]^{(\geqslant 3)} \cup \{A \in [n]^{(2)} : 1 \in A\}$ *is the unique* (*up to reordering the coordinates*) *family of size* $\sum_{k=3}^{n} \binom{n}{k} + n - 1$ *maximizing the number of intersecting subfamilies of size s,*
- $[n]^{(\geqslant 3)} \cup \{A \in [n]^{(2)} : n \notin A\}$ *is the unique* (*up to reordering the coordinates*) *family of size* $\sum_{k=3}^{n} \binom{n}{k} + \binom{n-1}{2}$ *maximizing the number of intersecting subfamilies of size s.*

Clearly these families are not nested, even with a reordering of the coordinates.

We can think of forming $\mathcal{A}(p)$ by first choosing a random variable $s \sim \text{Binom}(|\mathcal{A}|, p)$ and then choosing s sets uniformly at random from $\mathcal{A}$. Hence Theorem 1.1 shows that these families are the unique most probably intersecting families for these two specific sizes for any $0 < p < 1$.

If $p \ll 2^{-n}$, then the most likely scenario is that $\mathcal{A}_p$ is empty and the next most likely is that it consists of a single set. In each case the subfamily $\mathcal{A}_p$ is trivially intersecting regardless of our choice of the original family $\mathcal{A}$ (assuming, of course, that $\emptyset \notin \mathcal{A}$). The next most likely case is that there are exactly two sets in $\mathcal{A}_p$; this is far more likely than there being more than two sets. Hence, for very small p, proving Theorem 1.1 for the case $s = 2$ would give a counterexample to the conjecture. This was done by Ahlswede [1]. (Independently and slightly earlier, Frankl [3] proved the same result but without the necessary uniqueness.)

Our proof of Theorem 1.1 consists of two main steps summarized by the following two theorems.

Theorem 1.2. *Suppose that $n \geqslant 4$, $N \in \mathbb{N}$ and $2 \leqslant s \leqslant 2^{n-1}$. If*

$$n = 2t \text{ is even and } N > 2^{n-1} + \tfrac{1}{2}\binom{n}{t} - t$$

or

$$n = 2t + 1 \text{ is odd and } N > 2^{n-1} + \binom{n-1}{t-1} - t - 1,$$

then any family $\mathcal{A} \subset \mathcal{P}([n])$ of size N containing the maximal number of intersecting subfamilies of size s is of the form $[n]^{(\geqslant r+1)} \cup \mathcal{B}$, where $\mathcal{B} \subset [n]^{(r)}$ and r satisfies

$$\sum_{k=r+1}^{n} \binom{n}{k} \leqslant N < \sum_{k=r}^{n} \binom{n}{k}.$$

In [5] it is shown that there exists *some* family of this form maximizing the number of intersecting subfamilies of size s. Theorem 1.2 strengthens this result by showing that *all* the optimal families are of this form. This result may be of interest in its own right.

As we remarked above, the extremal families for $s = 2$ have been widely studied. However, it does not appear to have been proved, even in this case, that every extremal family must have the above form.

As $\mathcal{P}([n])$ contains many different intersecting families of order 2^{n-1}, we trivially require $N > 2^{n-1}$ in Theorem 1.2. In fact, it is easy to see that a larger lower bound on N is actually required. Indeed, if we take *any* maximal intersecting family $\mathcal{A}_0$ and form the family $\mathcal{A}$ by adding a maximal set A not in $\mathcal{A}_0$, then this new family is extremal for all s, since A and A^c are the only pair of disjoint sets in $\mathcal{A}$. In fact, the bound stated in Theorem 1.2 is tight: in Theorem 5.4 we construct, for all appropriate values of N, extremal families which are not of the desired form.

The final step is the following theorem, which is at the heart of the proof.

Theorem 1.3. *Suppose that $n \geqslant 21$. Let $2 \leqslant s \leqslant 2^{n-1}$ and $0 \leqslant i \leqslant \binom{n}{2}$. Suppose that $\mathcal{A} \subset \mathcal{P}([n])$ is any family of size $\sum_{k=3}^{n} \binom{n}{k} + i$ of the form $[n]^{(\geqslant 3)} \cup \mathcal{B}$ with $\mathcal{B} \subset [n]^{(2)}$, and that, subject to these conditions, $\mathcal{A}$ contains the maximal number of intersecting subfamilies of size s. Then $\mathcal{B}$ is a family of size i contained in $[n]^{(2)}$ that contains the maximal number of intersecting pairs.*

Note that we could rephrase the theorem to say that the family $\mathcal{A}$ that maximizes the number of intersecting subfamilies of size s necessarily also maximizes the number of intersecting pairs. This is clearly equivalent as each set in $\mathcal{B}$ intersects the same number of sets in $[n]^{(\geqslant 3)} = \mathcal{A} \setminus \mathcal{B}$.

Given Theorems 1.2 and 1.3, it is easy to prove Theorem 1.1.

Proof of Theorem 1.1. First suppose that $\mathcal{A}$ is a family of size $\sum_{k=3}^{n} \binom{n}{k} + n - 1$ maximizing the number of intersecting subfamilies of size s. Theorem 1.2 tells us that $\mathcal{A} = [n]^{(\geqslant 3)} \cup \mathcal{B}$ for some $\mathcal{B} \subset [n]^{(2)}$. Clearly we must have $|\mathcal{B}| = n - 1$. Now Theorem 1.3 tells us that $\mathcal{B}$ contains the maximal number of intersecting pairs over all families in $[n]^{(2)}$ of size $n - 1$. It is obvious that $\mathcal{B} = \{B \in [n]^{(2)} : 1 \in B\}$ maximizes the number of intersecting pairs, as all pairs intersect, and, since $|\mathcal{B}| > 3$, that it is the unique (up to reordering coordinates) family that does. Hence, in this case $\mathcal{A}$ must have the required form.

For the second case, suppose that $\mathcal{A}$ is a family of size $\sum_{k=3}^{n} \binom{n}{k} + \binom{n-1}{2}$ maximizing the number of intersecting subfamilies of size s. As above, we see that $\mathcal{A} = [n]^{(\geqslant 3)} \cup \mathcal{B}$ for some $\mathcal{B} \subset [n]^{(2)}$ and that $\mathcal{B}$ contains the maximal number of intersecting pairs over all families in $[n]^{(2)}$ of size $\binom{n-1}{2}$. Again, the extremal family $\mathcal{B}$ is unique up to reordering the coordinates: it consists of all the 2-sets not containing n. This is a little less obvious but follows from the result for $i = n - 1$ above. Indeed, the family $\mathcal{B}$ containing the most intersecting pairs minimizes the number of intersecting pairs with one element in $\mathcal{B}$ and one element in $\mathcal{B}^c$. Thus it also maximizes the number of intersecting pairs in $\mathcal{B}^c$. By the above, $\mathcal{B}^c$ is $\{A : 1 \in A\}$ and the result follows (after a reordering of the coordinates). □

In fact, the extremal families $\mathcal{B}$ have been precisely determined. Suppose $i = \binom{a}{2} + b$ with $0 \leqslant b < a$. The *quasi-complete graph* of order n with i edges is the graph formed by taking a complete graph on a vertices, adding a single vertex joined to b of the vertices of the complete graph and adding $n - a - 1$ isolated vertices. A *quasi-star* is the complement of a quasi-complete graph.

Ahlswede and Katona [2] showed that the families of 2-sets (graphs) with the most intersecting pairs (adjacent edges) are either quasi-complete graphs or quasi-stars. Moreover, they showed that there exists some non-negative integer R (depending on n) such that for $i < \frac{1}{2}\binom{n}{2} - R$ and for $\frac{1}{2}\binom{n}{2} \leqslant i \leqslant \frac{1}{2}\binom{n}{2} + R$ the extremal family is a quasi-star, while for all other values of i the extremal family is a quasi-complete graph. Wagner and Wang [6] extended this by finding the value of R explicitly and showing that it is non-zero for a proportion $\sqrt{2} - 1$ of numbers n. Combining Theorem 1.3 with these results we see that the extremal families even for N in this range are surprisingly complicated: for many values of n (*i.e.*, those for which $R \neq 0$) the extremal families can switch between the two classes three times just in this single layer.

Layout of the paper. In Section 2 we define the notation we shall use and recall the definitions and some of the properties of the compressions that we use. In Section 3 we prove Theorem 1.2 for the cases $N \geqslant \sum_{k=\lceil n/2 \rceil - 1}^{n} \binom{n}{k}$, which is sufficient (in combination with Theorem 1.3) to prove Theorem 1.1. In Section 4 we prove Theorem 1.3. In Section 5 we prove the remaining cases of Theorem 1.2 and give the constructions showing that the lower bound on N in Theorem 1.2 is tight. We conclude the paper in Section 6 with a discussion of some open problems.

2. Notation and preliminaries

Most of the notation we use is standard. We write $[n]$ for the set $\{1, 2, \ldots, n\}$ and $[m, n]$ for the set $\{m, m+1, \ldots, n\}$. For any r we use $[n]^{(r)}$ to denote the set of subsets of $[n]$ of size r, and $[n]^{(\geqslant r)}$ to denote the set of subsets of $[n]$ of size at least r.

For any family $\mathcal{A}$ we let $\mathfrak{I}^{(s)}(\mathcal{A})$ denote the collection of intersecting subfamilies of $\mathcal{A}$ of size s. For clarity, when s is clear from the context we suppress the superscript.

In much of this paper we shall be aiming to change or *compress* a family $\mathcal{A}$ into a nice form without decreasing the number of intersecting subfamilies of a given size.

We use three types of compression. The first is a very simple operation called an *up-set-compression*. We replace a set $A \in \mathcal{A}$ by a set $A' \supset A$ with $A' \notin \mathcal{A}$. Obviously this preserves the size of $\mathcal{A}$ and does not decrease the number of intersecting subfamilies.

The second operation we use is a very standard compression called an *ij-compression*. We take each set A in $\mathcal{A}$ and, if $i \notin A$ and $j \in A$, we replace A by the set $A \cup \{i\} \setminus \{j\}$ *provided* that this set is not already in $\mathcal{A}$. We note that these compressions do not change the size of any set in $\mathcal{A}$.

Again it is easy to see that this preserves the size of $\mathcal{A}$. This time, it is not obvious that the compression does not decrease the number of intersecting subfamilies. It is, however, proved in [5].

We will generally be applying these compressions when $i < j$ and we call such a compression a *left-compression*.

The final operation we use is the (U, v, f)-compression recently introduced in [5]. Suppose that $U \subset [n]$ has even size, that $f : U \to U$ is a permutation of order 2 with no fixed point, and that $v \in [n] \setminus U$. We move each set $A \in \mathcal{A}$ with $v \notin A$ to $(A \setminus U \cup \{v\}) \cup f(A \cap U)$ unless this set is already in $\mathcal{A}$. Informally, we add v and swap the points inside U. Again it is clear that this does not change the size of $\mathcal{A}$. Note also that every set moved by this compression contains v after the move.

We shall use the following key property of these (U, v, f)-compressions (proved in [5]). For any such compression C there exists an injection $\widehat{C}$ from $\mathfrak{I}^{(s)}(\mathcal{A})$ to $\mathfrak{I}^{(s)}(C(\mathcal{A}))$ and so, in particular, the number of intersecting subfamilies of any given order does not decrease. The only property of $\widehat{C}$ that we shall use is that $\widehat{C}(\mathcal{B}) \in \mathfrak{I}(C(\mathcal{A}))$ is formed from $\mathcal{B} \in \mathfrak{I}(\mathcal{A})$ by sending each set $A \in \mathcal{B}$ to either A or $C(A)$. We remark that constructing the injection $\widehat{C}$ is non-trivial.

3. Proof of Theorem 1.2 for $N \geqslant \sum_{k=\lceil n/2 \rceil - 1}^{n} \binom{n}{k}$

In this section we prove Theorem 1.2 for all the cases where $N \geqslant \sum_{k=\lceil n/2 \rceil - 1}^{n} \binom{n}{k}$: that is, the N for which our putative extremal family would contain all of the first layer below the middle. This is sufficient for our main result (Theorem 1.1). For completeness, we prove the remaining cases in Section 5.

Define $r = r(N, n)$ to be the unique number r satisfying

$$\sum_{k=r+1}^{n} \binom{n}{k} \leqslant N < \sum_{k=r}^{n} \binom{n}{k}.$$

Thus the bound for N above corresponds to $r < n/2 - 1$.

We start by showing that if $\mathcal{A}$ has a particularly nice form then there is a (U, v, f)-compression that strictly increases the number of intersecting subfamilies of size s.

Lemma 3.1. *Let* $2 \leqslant s \leqslant 2^{n-1}$ *and* $\ell < \frac{n}{2} - 1$. *Suppose that* $\mathcal{A}$ *satisfies* $[n]^{(>\ell+1)} \subset \mathcal{A}$, $[n - \ell, n] \notin \mathcal{A}$ *and* $[\ell] \in \mathcal{A}$. *Then there is a* (U, v, f)-*compression* C *such that* $|\mathfrak{I}^{(s)}(C(\mathcal{A}))| > |\mathfrak{I}^{(s)}(\mathcal{A})|$.

Proof. Choose C to be any (U, v, f)-compression with $v = n$ that moves $[\ell]$ to $[n - \ell, n]$, and let $\mathcal{C} = C(\mathcal{A})$ be the resulting family. We construct a family in $\mathfrak{I}^{(s)}(\mathcal{C})$ that is not the image of any family in $\mathfrak{I}^{(s)}(\mathcal{A})$ under the injection $\widehat{C}$.

Consider the family

$$\mathcal{D} = \{[n - \ell, n]\} \cup [n]^{(\geqslant n - \ell - 1)} \setminus \{[n - \ell - 1]\}.$$

This is intersecting, since $\ell < n/2 - 1$ and so $n - \ell - 1 > n/2$. Thus it extends to a maximal intersecting family $\mathcal{D}'$ of size 2^{n-1} in $\mathcal{P}([n])$. Since $\mathcal{D}'$ contains all sets of size at least $n - \ell - 1$ except $[n - \ell - 1]$ and is intersecting, $\mathcal{D}'$ contains no set of size less than or equal to $\ell + 1$ except $[n - \ell, n]$. By hypothesis, $\mathcal{A}$ contains all of $[n]^{(>\ell+1)}$ and thus so does $\mathcal{C}$. Moreover, $[\ell] \in \mathcal{A}$ is moved to $[n - \ell, n]$ so $[n - \ell, n] \in \mathcal{C}$. Hence $\mathcal{D}' \subset \mathcal{C}$. Also, since $|[\ell + 1, n - 1]| = n - \ell - 1$, we have $[\ell + 1, n - 1] \in \mathcal{D}'$. Let $\mathcal{D}''$ be any subfamily of size s of $\mathcal{D}'$ containing both $[n - \ell, n]$ and $[\ell + 1, n - 1]$. Note that $\mathcal{D}'' \in \mathfrak{I}(\mathcal{C})$.

Suppose that there is an intersecting subfamily $\mathcal{B}$ of $\mathcal{A}$ with $\widehat{C}(\mathcal{B}) = \mathcal{D}''$. Recall that $\widehat{C}(\mathcal{B})$ is formed from $\mathcal{B}$ by sending each set $A \in \mathcal{B}$ to either A or $C(A)$. Now $[n - \ell, n] \in \mathcal{D}''$ but $[n - \ell, n] \notin \mathcal{A}$, so $[n - \ell, n] \notin \mathcal{B}$. Hence $[n - \ell, n]$ must have come from $[\ell] \in \mathcal{B}$. Also, $[\ell + 1, n - 1] \in \mathcal{D}''$ and, since $n \notin [\ell + 1, n - 1]$, this set has a unique pre-image under C, namely the set $[\ell + 1, n - 1]$ itself. Therefore $[\ell + 1, n - 1] \in \mathcal{B}$. But we also have $[\ell] \in \mathcal{B}$, contradicting the fact that $\mathcal{B}$ is intersecting.

We conclude that $\mathcal{D}''$ is not the image under $\widehat{C}$ of any family in $\mathfrak{I}^{(s)}(\mathcal{A})$. Hence $|\mathfrak{I}^{(s)}(\mathcal{C})| > |\mathfrak{I}^{(s)}(\mathcal{A})|$. $\qquad\square$

In Section 5 we slightly strengthen this result, proving that with some extra conditions it holds for $\ell = \lceil n/2 \rceil - 1$.

Corollary 3.2. *Let* $n, N \in \mathbb{N}$ *with* $r = r(N, n) < n/2$ *and* $2 \leqslant s \leqslant 2^{n-1}$. *Suppose that* $[n]^{(\geqslant r+1)} \subset \mathcal{A}$ *and that* $\mathcal{A}$ *contains a set of size strictly less than* r. *Then there is a family of size N that contains strictly more intersecting subfamilies of size s than does* $\mathcal{A}$.

Proof. By the definition of r we see that $\mathcal{A}$ does not contain all of $[n]^{(r)}$. Hence by applying left-compressions we can ensure that $[n - r + 1, n] \notin \mathcal{A}$. Also, since $\mathcal{A}$ contains some set of size at most $r - 1$, by applying up-set-compressions and left-compressions we can ensure that $[r - 1] \in \mathcal{A}$; it is easy to check that we can do this without putting $[n - r + 1, n]$ into $\mathcal{A}$. Thus Lemma 3.1 applies with $\ell = r - 1$. □

Lemma 3.3. *Suppose that* $n, N \in \mathbb{N}$ *with* $r = r(N, n) < n/2 - 1$, *that* $[n]^{(\geqslant r+1)} \not\subset \mathcal{A}$, *and that* $2 \leqslant s \leqslant 2^{n-1}$. *Then there is a family that contains strictly more intersecting subfamilies of size s than does* $\mathcal{A}$.

Proof. We aim to compress the family until it contains nearly all of $[n]^{(\geqslant r+1)}$. We then apply one more compression and use Lemma 3.1 to show strict inequality for this final compression. We need to be careful that the earlier compressions do not 'accidentally' put all sets in $[n]^{(\geqslant r+1)}$ into our family since then we would not necessarily obtain strict inequality when applying the final compression.

We construct a sequence of families $\mathcal{A} = \mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k$, with $[n]^{(\geqslant r+1)} \not\subset \mathcal{A}_i$ for any i, by applying at each stage any 'allowed' up-set-compression, left-compression or (U, v, f)-compression – that is, one which does not result in our family containing the whole of $[n]^{(\geqslant r+1)}$. We finish with a family $\mathcal{A}_k$ that is unchanged by any compression C with $[n]^{\geqslant (r+1)} \not\subset C(\mathcal{A}_k)$. Note that $\mathcal{A}_k$ is left-compressed since ij-compressions do not change the size of any set. Also, by considering up-set-compressions, we see that $\mathcal{A}^+ = \mathcal{A}_k \cap [n]^{(\geqslant r+1)}$ is an up-set and $\mathcal{A}^- = \mathcal{A}_k \cap [n]^{(\leqslant r)}$ is an up-set when viewed as a subset of $[n]^{(\leqslant r)}$. Obviously $\mathcal{A}^+ \neq [n]^{(\geqslant r+1)}$ and so $\mathcal{A}^-$ is non-empty.

We claim that $\mathcal{A}^+ = [n]^{(\geqslant r+1)} \setminus \{[n - r, n]\}$. If only one set from $[n]^{(\geqslant r+1)}$ is missing from $\mathcal{A}^+$ then it must be $[n - r, n]$. Thus we may assume for a contradiction that at least two of the sets in $[n]^{(\geqslant r+1)}$ are missing from $\mathcal{A}^+$. Since $\mathcal{A}^+$ is a left-compressed up-set we see that these missing sets must include $[n - r, n]$ and $\{n - r - 1\} \cup [n - r + 1, n]$. Similarly, as $\mathcal{A}^-$ is a non-empty left-compressed 'up-set', we see that $[r] \in \mathcal{A}^-$. We have now shown that $[r] \in \mathcal{A}_k$, that $[n - r, n] \notin \mathcal{A}_k$ and that $\{n - r - 1\} \cup [n - r + 1, n] \notin \mathcal{A}_k$. The upper bound on r implies that the sets $[r]$ and $[n - r, n]$ are disjoint. Hence we can map the former to the latter using a (U, v, f)-compression with $v = n - r$. This does not add the set $\{n - r - 1\} \cup [n - r + 1, n]$ since all sets added by such a compression contain $v = n - r$. Hence this is an allowed (U, v, f)-compression which contradicts the definition of $\mathcal{A}_k$.

So $\mathcal{A}_k$ contains all of $[n]^{(\geqslant r+1)}$ except for the set $[n - r, n]$ and, as before, it must contain $[r]$. Hence Lemma 3.1 applies with $\ell = r$. □

This essentially completes the proof of Theorem 1.2 for $r < \frac{n}{2} - 1$. Indeed, by Lemma 3.3, $[n]^{(\geqslant r+1)} \subset \mathcal{A}$ and so, by Corollary 3.2, $\mathcal{A}$ has the required form.

The only remaining cases are $\frac{n}{2} - 1 \leqslant r \leqslant \frac{n}{2}$. We deal with these cases in Section 5.

4. Proof of Theorem 1.3

Fix s and, as usual, let $\mathfrak{I} = \mathfrak{I}^{(s)}$ denote the collection of intersecting subfamilies of $\mathcal{A}$ of size s. For $\mathcal{B} \subset \mathcal{A} \cap [n]^{(2)}$ let

$$\mathfrak{I}_{\mathcal{B}} = \{\mathcal{E} \in \mathfrak{I} : \mathcal{E} \cap [n]^{(2)} = \mathcal{B}\}.$$

We see that $\mathfrak{I}$ is the disjoint union of the sets $\mathfrak{I}_{\mathcal{B}}$ over all collections $\mathcal{B}$ of 2-sets in $\mathcal{A}$. Moreover, $\mathfrak{I}_{\mathcal{B}}$ is empty unless $\mathcal{B}$ is intersecting. Since all sets in $\mathcal{B}$ have size 2, the structure of these intersecting families is simple. Indeed, for all $0 \leqslant r \leqslant n - 1$ except $r = 3$, there is a unique (up to reordering the coordinates) intersecting family of size r in $[n]^{(2)}$, namely the star $\mathcal{S}_r$ consisting of the sets $\{1t\}$ for $2 \leqslant t \leqslant r + 1$. Trivially, for $n \geqslant 4$ there is no intersecting family of size greater than $n - 1$. For $r = 3$ there are two intersecting families (again up to reordering the coordinates), namely $\mathcal{S}_3 = \{12, 13, 14\}$ and $\mathcal{T} = \{12, 13, 23\}$, which we shall call the star and the triangle respectively.

Since, by hypothesis, we know that $\mathcal{A}$ contains all sets of size at least 3 and no sets of size 1, we have

$$\mathfrak{I}_{\mathcal{B}} = \{\mathcal{E} \subset \mathcal{P}([n]^{(\geqslant 2)}) : \mathcal{E} \cap [n]^{(2)} = \mathcal{B}, \mathcal{E} \text{ intersecting}\}.$$

Hence the cardinality of $\mathfrak{I}_{\mathcal{B}}$ depends only on which of $\mathcal{S}_0, \mathcal{S}_1, \ldots \mathcal{S}_{n-1}, \mathcal{T}$ is isomorphic to $\mathcal{B}$. Let

$$\mathfrak{I}_r = \{\mathcal{E} \subset \mathcal{P}([n]^{(\geqslant 2)}) : \mathcal{E} \cap [n]^{(2)} = \mathcal{S}_r, \mathcal{E} \text{ intersecting}\}$$

and

$$\mathfrak{I}_{\mathcal{T}} = \{\mathcal{E} \subset \mathcal{P}([n]^{(\geqslant 2)}) : \mathcal{E} \cap [n]^{(2)} = \mathcal{T}, \mathcal{E} \text{ intersecting}\}.$$

Let a_r be the number of intersecting subfamilies of $\mathcal{A} \cap [n]^{(2)}$ isomorphic to $\mathcal{S}_r$ and let b be the number isomorphic to $\mathcal{T}$. Then

$$|\mathfrak{I}| = \sum_{r=0}^{n-1} a_r |\mathfrak{I}_r| + b |\mathfrak{I}_{\mathcal{T}}|.$$

Obviously $a_0 = 1$ and $a_1 = |\mathcal{A} \cap [n]^{(2)}| = i$ so the first two terms of the sum are independent of the collection $\mathcal{A}$. Trivially we have $a_r \leqslant n\binom{n-1}{r}$ and $b \leqslant \binom{n}{3}$. If we compare $\mathfrak{I}_r$ and $\mathfrak{I}_{r-1}$, we see that a family in $\mathfrak{I}_r$ has two extra restrictions: it must contain the set $\{1, r+1\}$ (which gives us one fewer set to place) and each set must intersect $\{1, r+1\}$ (which places an extra restriction on where these other sets can lie). Thus we might expect $\mathfrak{I}_{r-1}$ to be much larger than $\mathfrak{I}_r$. That is precisely what Lemma 4.2 will show. First we need the following simple result.

Lemma 4.1. $|\mathfrak{I}_{\mathcal{T}}| \leqslant |\mathfrak{I}_3|$.

Proof. There is a unique maximal intersecting family containing $\mathcal{T}$ and thus every intersecting family containing $\mathcal{T}$ is contained in this unique maximal family. Hence the number of intersecting families of size s containing $\mathcal{T}$ is the smallest it can possibly be, namely $\binom{2^{n-1}-3}{s-3}$. $\qquad \square$

We give one definition that will be useful in the proof.

Definition. Suppose that $\mathcal{E}, \mathcal{F}$ are families in $\mathcal{P}([n])$. We say they are *cross-intersecting* (or that they *cross-intersect*) if for every $E \in \mathcal{E}$ and $F \in \mathcal{F}$ we have $E \cap F \neq \emptyset$.

Lemma 4.2. *If* $r > 3$ *then*

$$|\mathfrak{I}_{r-1}| \geqslant \left(2^{n-r-2} - \frac{n-r}{2} \right) |\mathfrak{I}_r|,$$

and if $r = 3$ *then*

$$|\mathfrak{I}_{r-1}| \geqslant \left(2^{n-5} - \frac{n-1}{2} \right) |\mathfrak{I}_r|.$$

Proof. For $r > 3$ we construct a mapping $\Phi \colon [r+2,n]^{(\geqslant 2)} \times \mathfrak{I}_r \to \mathfrak{I}_{r-1}$ under which every family in $\mathfrak{I}_{r-1}$ has at most two pre-images. In the case $r = 3$ we instead construct $\Phi \colon [r+2,n]^{(\geqslant 2)} \times \mathfrak{I}_r \to \mathfrak{I}_{r-1} \cup \mathfrak{I}_{\mathcal{T}}$. Recalling that $|\mathfrak{I}_{\mathcal{T}}| \leqslant |\mathfrak{I}_3|$, this suffices to prove the lemma.

Throughout the proof we write X^c to denote the complement of the set X relative to $[r+2,n]$: that is, for $X \subset [r+2,n]$ we write $X^c = [r+2,n] \setminus X$.

Suppose that $\mathcal{E} \in \mathfrak{I}_r$ and $U \in [r+2,n]^{(\geqslant 2)}$. Let $U' = \{1, r+1\} \cup U$. First we tweak $\mathcal{E}$ slightly to make sure that it contains U'. If $U' \in \mathcal{E}$ let $\bar{\mathcal{E}} = \mathcal{E}$; otherwise let $\bar{\mathcal{E}} = \mathcal{E} \setminus \{\{1, r+1\}\} \cup \{U'\}$. Note that the new family $\bar{\mathcal{E}}$ is still intersecting since $\{1, r+1\} \subset U'$.

We split $\bar{\mathcal{E}}$ into pieces as follows:

$$\mathcal{E}_0 = \{E \in \bar{\mathcal{E}} : 1 \in E, E \cap [2,r] \neq \emptyset\},$$
$$\mathcal{E}_1 = \{E \in \bar{\mathcal{E}} : 1 \in E, E \cap [2,r+1] = \emptyset\},$$
$$\mathcal{E}_2 = \{E \in \bar{\mathcal{E}} : 1 \notin E\},$$
$$\mathcal{E}_3 = \{E \in \bar{\mathcal{E}} : 1, r+1 \in E, E \cap [2,r] = \emptyset\}.$$

Clearly $\mathcal{E}_2 = \{E \in \mathcal{E} : 1 \notin E\}$. As $\mathcal{E} \in \mathfrak{I}_r$ we know that $\mathcal{E}$ is intersecting and $\{1, j\} \in \mathcal{E}$ for $2 \leqslant j \leqslant r+1$. Hence $\mathcal{E}_2 = \{E \in \bar{\mathcal{E}} : E \cap [r+1] = [2, r+1]\}$.

We define $\mathcal{E}'_1, \mathcal{E}'_2, \mathcal{E}'_3$ to be the restrictions of $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$ to $[r+2,n]$: that is, $\mathcal{E}'_i = \{E \cap [r+2,n] : E \in \mathcal{E}_i\}$ for $i = 1, 2, 3$. Since the intersection of a set in $\mathcal{E}_i$ ($i = 1, 2, 3$) with $[1, r+1]$ depends only on i, we see that the sets $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$ are determined by $\mathcal{E}'_1, \mathcal{E}'_2, \mathcal{E}'_3$ respectively.

We make a couple of remarks about this partition that will be helpful later in the proof. First, $U' \in \mathcal{E}_3$ and so $U \in \mathcal{E}'_3$. Secondly, $\mathcal{E}'_1$ and $\mathcal{E}'_2$ are cross-intersecting.

To define our new intersecting family $\mathcal{F} = \Phi(U, \mathcal{E})$ we split into two cases according to whether $U^c \cap E \neq \emptyset$ for all $E \in \mathcal{E}'_1$. Note that if this condition does not hold then U meets every element of $\mathcal{E}'_2$. Indeed, suppose $F \in \mathcal{E}'_2$ with $F \cap U = \emptyset$ and $E \in \mathcal{E}'_1$ with $E \cap U^c = \emptyset$. Then $F \subset U^c$ and $E \subset U$ so $E \cap F = \emptyset$, which contradicts the cross-intersection property observed above.

Case 1: $U^c \cap E \neq \emptyset$ for all $E \in \mathcal{E}'_1$. This is the simpler case. Here, starting from $\bar{\mathcal{E}}$, we replace each set $X \in \bar{\mathcal{E}}$ satisfying $\{1, r+1\} \subset X \subset U'$ by its complement. Formally, let $\mathcal{F}_i = \mathcal{E}_i$ for

$i = 0, 1, 2$, let

$$\mathcal{F}_3 = \{\{1, r+1\} \cup E : E \in \mathcal{E}'_3, E \cap U^c \neq \emptyset\},$$

and let

$$\mathcal{F}_4 = \{[2, r] \cup E^c : E \in \mathcal{E}'_3, E \subseteq U\}.$$

Set $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{F}_4$.

The families $\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_4$ are pairwise disjoint and there is an obvious bijection from $\mathcal{E}_3$ to $\mathcal{F}_3 \cup \mathcal{F}_4$. Hence $|\mathcal{F}| = |\mathcal{E}|$.

Moreover, this new family $\mathcal{F}$ is intersecting: obviously $\mathcal{F}_0 \cup \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3 \subset \mathcal{E}$ and so is intersecting, and $\mathcal{F}_4$ is an intersecting family, so we only need to check that $\mathcal{F}_4$ cross-intersects each of the other $\mathcal{F}_i$. Trivially $\mathcal{F}_4$ cross-intersects $\mathcal{F}_0$ and $\mathcal{F}_2$. We see that $\mathcal{F}_3$ and $\mathcal{F}_4$ cross-intersect as U^c intersects every set in $\mathcal{F}_3$ and is contained in every set in $\mathcal{F}_4$. Finally, $\mathcal{F}_1$ and $\mathcal{F}_4$ cross-intersect because we are assuming that U^c intersects everything in $\mathcal{E}'_1$.

Next we show that if $r \geqslant 4$ then $\mathcal{F} \cap [n]^{(\leqslant 2)}$ is $\mathcal{S}_{r-1}$, and if $r = 3$ then $\mathcal{F} \cap [n]^{(\leqslant 2)}$ is either $\mathcal{S}_{r-1}$ or $\mathcal{T}$. Since $\mathcal{E} \cap [n]^{(\leqslant 2)} = \mathcal{S}_r$, we see that $\bar{\mathcal{E}} \cap [n]^{(\leqslant 2)}$ is either $\mathcal{S}_r$ or $\mathcal{S}_{r-1}$. Clearly $\mathcal{S}_{r-1} \subset \mathcal{E}_0$. If $\{1, r+1\} \in \bar{\mathcal{E}}$ then it is in $\mathcal{E}_3$ but not $\mathcal{F}_3$. Hence $(\mathcal{F}_0 \cup \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3) \cap [n]^{(\leqslant 2)} = \mathcal{S}_{r-1}$. Finally, if $r \geqslant 4$ then $\mathcal{F}_4 \cap [n]^{(\leqslant 2)} = \emptyset$; if $r = 3$ then $\mathcal{F}_4 \cap [n]^{(\leqslant 2)}$ is either empty or $\{\{2, 3\}\}$. Thus we have shown that if $r \geqslant 4$ then $\mathcal{F} \in \mathfrak{I}_{r-1}$ and if $r = 3$ then $\mathcal{F} \in \mathfrak{I}_{r-1} \cup \mathfrak{I}_{\mathcal{T}}$.

Finally, *if we know that $\mathcal{F}$ comes from this case* then we can reconstruct $\bar{\mathcal{E}}$. Indeed, given $\mathcal{F}$, set $\mathcal{F}_4 = \{F \in \mathcal{F} : F \cap [1, r+1] = [2, r]\}$. Then form $\bar{\mathcal{E}}$ from $\mathcal{F}$ by replacing each set in $\mathcal{F}_4$ by its complement. We also know U since $[2, r] \cup U^c$ is the unique minimal element of $\mathcal{F}_4$. Once we know $\bar{\mathcal{E}}$ and U, it easy to reconstruct $\mathcal{E}$.

Case 2: $U \cap E \neq \emptyset$ for all $E \in \mathcal{E}'_2$. This time the construction is a little more complicated. We define

$$\begin{aligned}
\mathcal{F}_0 &= \mathcal{E}_0, \\
\mathcal{F}_1 &= \{\{1, r+1\} \cup E : E \in \mathcal{E}'_1\}, \\
\mathcal{F}_2 &= \{[2, r] \cup E : E \in \mathcal{E}'_2\}, \\
\mathcal{F}_3 &= \{\{1\} \cup E : E \in \mathcal{E}'_3, U \subseteq E\}, \\
\mathcal{F}_4 &= \{[2, r+1] \cup E^c : E \in \mathcal{E}'_3, E^c \cap U \neq \emptyset\},
\end{aligned}$$

and, as before, set $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{F}_4$.

Again $\mathcal{F}_0 = \mathcal{E}_0$ but this time $\mathcal{F}_1 \neq \mathcal{E}_1$ and $\mathcal{F}_2 \neq \mathcal{E}_2$. However, there are bijections between $\mathcal{F}_1$ and $\mathcal{E}_1$, and between $\mathcal{F}_2$ and $\mathcal{E}_2$. As before there is a bijection between $\mathcal{F}_3 \cup \mathcal{F}_4$ and $\mathcal{E}_3$. Hence $|\mathcal{F}| = |\mathcal{E}|$.

Again $\mathcal{F}$ is intersecting. Indeed, each $\mathcal{F}_i$ is trivially intersecting, $\mathcal{F}_0$ is cross-intersecting with each of the others, and each of the pairs $(\mathcal{F}_1, \mathcal{F}_3)$, $(\mathcal{F}_1, \mathcal{F}_4)$ and $(\mathcal{F}_2, \mathcal{F}_4)$ is trivially cross-intersecting. We see that $\mathcal{F}_3$ and $\mathcal{F}_4$ cross-intersect as U is contained in every set in $\mathcal{F}_3$ and intersects every set in $\mathcal{F}_4$. Also, $\mathcal{F}_1$ and $\mathcal{F}_2$ cross-intersect because $\mathcal{E}'_1$ and $\mathcal{E}'_2$ are cross-intersecting. Finally, $\mathcal{F}_2$ and $\mathcal{F}_3$ cross-intersect because we are assuming that U intersects everything in $\mathcal{E}'_2$.

Next we show that if $r \geqslant 4$ then $\mathcal{F} \cap [n]^{(\leqslant 2)}$ is $\mathcal{S}_{r-1}$, and if $r = 3$ then $\mathcal{F} \cap [n]^{(\leqslant 2)}$ is either $\mathcal{S}_{r-1}$ or $\mathcal{T}$. We consider each $\mathcal{F}_i \cap [n]^{(\leqslant 2)}$ in turn. We have $\mathcal{F}_0 \cap [n]^{(\leqslant 2)} = \mathcal{E}_0 \cap [n]^{(\leqslant 2)} = \mathcal{S}_{r-1}$. As

$\{1\} \notin \mathcal{E}, \emptyset \notin \mathcal{E}_1'$ so $\mathcal{F}_1 \cap [n]^{(\leqslant 2)} = \emptyset$. If $r \geqslant 4$ then $\mathcal{F}_2 \cap [n]^{(\leqslant 2)} = \emptyset$; if $r = 3$ then $\mathcal{F}_2 \cap [n]^{(\leqslant 2)}$ is either empty or $\{\{2,3\}\}$. As $|U| \geqslant 2$, $\mathcal{F}_3 \cap [n]^{(\leqslant 2)} = \emptyset$. Finally, it is obvious that $\mathcal{F}_4 \cap [n]^{(\leqslant 2)} = \emptyset$. Again we have shown that if $r \geqslant 4$ then $\mathcal{F} \in \mathfrak{I}_{r-1}$ and if $r = 3$ then $\mathcal{F} \in \mathfrak{I}_{r-1} \cup \mathfrak{I}_{\mathcal{T}}$.

Now, given $\mathcal{F}$ we can determine $\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_4$ by considering intersections with $[1, r+1]$. Thus, if we knew we were in this case, we could reconstruct $\mathcal{E}$ and U. (This time $\{1\} \cup U$ is the unique minimal element of $\mathcal{F}_3$.)

However, we cannot (necessarily) tell from which case the family $\mathcal{F}$ came. Thus the function is not necessarily injective but each family in $\mathfrak{I}_{r-1}$ has at most two pre-images as required. $\qquad \square$

The rest of the proof is straightforward calculation. Recall that, by hypothesis, $n \geqslant 21$. It follows from Lemma 4.2 that $|\mathfrak{I}_3| \leqslant 2^{6-n}|\mathfrak{I}_2|$. It also follows that $|\mathfrak{I}_4| \leqslant 2^{7-n}|\mathfrak{I}_3|$, and that $|\mathfrak{I}_r| \leqslant |\mathfrak{I}_{r-1}|$ for $r \geqslant 5$. Thus, for all $r \geqslant 4$, $|\mathfrak{I}_r| \leqslant 2^{13-2n}|\mathfrak{I}_2|$. Furthermore, by Lemmas 4.1 and 4.2, $|\mathfrak{I}_{\mathcal{T}}| \leqslant 2^{6-n}|\mathfrak{I}_2|$. Recall $a_0 = 1$ and $a_1 = i$. Now

$$
\begin{aligned}
|\mathfrak{I}| &= \sum_{r=0}^{n-1} a_r |\mathfrak{I}_r| + b|\mathfrak{I}_{\mathcal{T}}| \\
&= |\mathfrak{I}_0| + i|\mathfrak{I}_1| + a_2|\mathfrak{I}_2| + \sum_{r=3}^{n-1} a_r|\mathfrak{I}_r| + b|\mathfrak{I}_{\mathcal{T}}| \\
&= |\mathfrak{I}_0| + i|\mathfrak{I}_1| + |\mathfrak{I}_2|\left(a_2 + \sum_{r=3}^{n-1} a_r \frac{|\mathfrak{I}_r|}{|\mathfrak{I}_2|} + b\frac{|\mathfrak{I}_{\mathcal{T}}|}{|\mathfrak{I}_2|} \right),
\end{aligned}
$$

and

$$
\sum_{r=3}^{n-1} a_r \frac{|\mathfrak{I}_r|}{|\mathfrak{I}_2|} + b\frac{|\mathfrak{I}_{\mathcal{T}}|}{|\mathfrak{I}_2|} \leqslant n\binom{n-1}{3}2^{6-n} + \binom{n}{3}2^{6-n} + \sum_{r=4}^{n-1} n\binom{n-1}{r}2^{13-2n}.
$$

It is easy to verify that the quantity on the right-hand side of the inequality is less than 1 for all $n \geqslant 21$.

This essentially completes the proof. Indeed, suppose $\mathcal{B}' \subset [n]^{(2)}$ is a family of size i with strictly more intersecting pairs than $\mathcal{B}$. Let a_r' and b' be the corresponding values for $\mathcal{B}'$. Then $a_0' = 1$, $a_1' = i$, and $a_2' \geqslant a_2 + 1$. Let $\mathfrak{I}'$ be the collection of intersecting families of size s in $[n]^{(\geqslant 3)} \cup \mathcal{B}'$. Using the above we have

$$
\begin{aligned}
|\mathfrak{I}'| &= \sum_{r=0}^{n-1} a_r'|\mathfrak{I}_r| + b'|\mathfrak{I}_{\mathcal{T}}| \\
&\geqslant |\mathfrak{I}_0| + i|\mathfrak{I}_1| + a_2'|\mathfrak{I}_2| \\
&\geqslant |\mathfrak{I}_0| + i|\mathfrak{I}_1| + (a_2 + 1)|\mathfrak{I}_2| \\
&> \sum_{r=0}^{n-1} a_r|\mathfrak{I}_r| + b|\mathfrak{I}_{\mathcal{T}}| \\
&= |\mathfrak{I}|,
\end{aligned}
$$

contradicting the maximality of $\mathcal{A}$. $\qquad \square$

5. The middle-layer cases of Theorem 1.2

In this section we conclude the proof of Theorem 1.2 by showing that it holds for $\frac{n}{2} - 1 \leqslant r \leqslant \frac{n}{2}$. We consider separately the cases of n even and n odd.

First we deal with some cases where N is not too close to the bound stated in Theorem 1.2. In each case we prove a slight variant on Lemma 3.1 and use it to deduce the result.

Case 1: $n = 2t$ and

$$\sum_{k=t}^{n} \binom{n}{k} \leqslant N < \sum_{k=t-1}^{n} \binom{n}{k}.$$

The bounds on N imply that $r = t - 1 = \frac{n}{2} - 1$. Also, the lower bound on N is equal to $2^{n-1} + \frac{1}{2} \binom{n}{t}$ so this covers nearly all of the remaining possibilities for N.

Lemma 5.1. *Let n be even, $2 \leqslant s \leqslant 2^{n-1}$ and $\ell = \frac{n}{2} - 1$. Suppose that $\mathcal{A}$ satisfies $[n]^{(\geqslant \ell + 1)} \setminus \{[n - \ell, n]\} \subset \mathcal{A}$, $[n - \ell, n] \notin \mathcal{A}$ and $[\ell] \in \mathcal{A}$. Then there is a (U, v, f)-compression C such that $|\mathfrak{I}^{(s)}(C(\mathcal{A}))| > |\mathfrak{I}^{(s)}(\mathcal{A})|$.*

Proof. Exactly as before, choose C to be any (U, v, f)-compression with $v = n$ that moves $[\ell]$ to $[n - \ell, n]$, and let $\mathcal{C} = C(\mathcal{A})$ be the resulting family. We construct a family in $\mathfrak{I}^{(s)}(\mathcal{C})$ that is not the image of any family in $\mathfrak{I}^{(s)}(\mathcal{A})$ under the injection $\widehat{C}$.

Consider the family

$$\mathcal{D} = \{[n - \ell, n], [n - \ell - 1, n - 1]\} \cup [n]^{(> \ell + 1)}.$$

Note that $|[n - \ell, n]| = |[n - \ell - 1, n - 1]| = \ell + 1 = \frac{n}{2}$ and so $\mathcal{D}$ is intersecting. Thus it extends to a maximal intersecting family $\mathcal{D}'$ of size 2^{n-1} in $\mathcal{P}([n])$. Since $\mathcal{D}'$ contains all sets of size at least $\ell + 2 = n/2 + 1$ and is intersecting, it contains no set of size less than or equal to ℓ. By hypothesis $\mathcal{A}$ contains all of $[n]^{(\geqslant \ell + 1)}$ except $[n - \ell, n]$ and thus so does $\mathcal{C}$. Moreover $[\ell] \in \mathcal{A}$ is moved to $[n - \ell, n]$ so $[n - \ell, n] \in \mathcal{C}$. Thus, in fact, $\mathcal{C}$ contains all of $[n]^{(\geqslant \ell + 1)}$. Hence $\mathcal{D}' \subset \mathcal{C}$.

Let $\mathcal{D}''$ be any subfamily of $\mathcal{D}'$ of size s containing $[n - \ell, n]$ and $[n - \ell - 1, n - 1]$. Note that $\mathcal{D}'' \in \mathfrak{I}(\mathcal{C})$. Exactly as before, we see that any intersecting family mapping to $\mathcal{D}''$ under $\widehat{C}$ would have to contain $[\ell]$ and $[n - \ell - 1, n - 1]$, which is not possible as $[n - \ell - 1, n - 1] = [\ell + 1, n - 1]$ is disjoint from $[\ell]$. ∎

Proof of Theorem 1.2 for this case. Using Lemma 5.1 instead of Lemma 3.1 we can prove a result analogous to Lemma 3.3. Combining this with Corollary 3.2 is sufficient to complete the proof in this case. ∎

Case 2: $n = 2t + 1$ and

$$\sum_{k=t+1}^{n} \binom{n}{k} + \binom{n-1}{t-1} \leqslant N < \sum_{k=t}^{n} \binom{n}{k}.$$

In this case the bounds on N imply that $r = t = \frac{n-1}{2}$. Also, the lower bound on N is equal to $2^{n-1} + \binom{n-1}{t-1}$.

Lemma 5.2. *Let n be odd, $2 \leqslant s \leqslant 2^{n-1}$ and $\ell = \frac{n-1}{2}$. Suppose that $\mathcal{A}$ satisfies $[n]^{(\geqslant \ell+1)} \setminus \{[n-\ell, n]\} \subset \mathcal{A}$ and $[n-\ell, n] \notin \mathcal{A}$, and that there exist $A, A' \in [n]^{(\ell)} \cap \mathcal{A}$ with $n \notin A$, $n \notin A'$ and $A \cap A' = \emptyset$. Then there is a (U, v, f)-compression C with $|\mathfrak{I}^{(s)}(C(\mathcal{A}))| > |\mathfrak{I}^{(s)}(\mathcal{A})|$.*

Proof. As A and A' are distinct we may assume without loss of generality that $A' \neq [\ell - 1]$. In particular this implies that $A' \cap [n - \ell, n] \neq \emptyset$.

This time we choose C to be any (U, v, f)-compression with $v = n$ that moves A to $[n - \ell, n]$. Let $\mathcal{C} = C(\mathcal{A})$ be the resulting family. We again construct a family in $\mathfrak{I}^{(s)}(\mathcal{C})$ that is not the image of any family in $\mathfrak{I}^{(s)}(\mathcal{A})$ under the injection $\widehat{C}$.

Let

$$\mathcal{D} = [n]^{(\geqslant \ell+1)} \setminus \{[n] \setminus A'\} \cup \{A'\}.$$

This is itself a maximal intersecting family. Since $C(A) = [n - \ell, n]$, we see that $\mathcal{C}$ contains all of $[n]^{(\geqslant \ell+1)}$. Further, since A is the only set in $\mathcal{A}$ of size ℓ that moves, we see that $A' \in \mathcal{C}$. Hence $\mathcal{D} \subset \mathcal{C}$.

Let $\mathcal{D}''$ be any subfamily of size s of $\mathcal{D}$ containing $[n - \ell, n]$ and A'. Then $\mathcal{D}'' \in \mathfrak{I}(\mathcal{C})$. Similarly to the previous cases, we see that any intersecting family mapping to $\mathcal{D}''$ under $\widehat{C}$ would have to contain A and A', which is not possible. $\qquad \square$

Proof of Theorem 1.2 for this case. Here we require a little more care. If $[n]^{(\geqslant r+1)} \subset \mathcal{A}$ then Corollary 3.2 tells us that $\mathcal{A}$ has the required form. Hence we may assume that $\mathcal{A}$ does not contain all of $[n]^{(\geqslant r+1)}$. We aim to compress $\mathcal{A}$ into a form where we can apply Lemma 5.2.

Exactly as in the proof of Lemma 3.3, we construct a sequence of families $\mathcal{A} = \mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k$ such that $[n]^{(\geqslant r+1)} \not\subset \mathcal{A}_i$ for any i, by applying at each stage an allowed compression. Let $\mathcal{A}' = \mathcal{A}_k$ be the final compressed family. As before, the only set of $[n]^{(\geqslant r+1)}$ not in $\mathcal{A}'$ is $[n - r, n]$.

Since $t = r$, the lower bound on N implies that $|\mathcal{A}' \cap [n]^{(\leqslant r)}| \geqslant \binom{n-1}{r-1} + 1$. We claim that, in fact, there are at least this many sets in the layer $[n]^{(r)}$.

If $\mathcal{A}'$ contains no set of size less than r then the claim holds trivially. Otherwise, $\mathcal{A}' \cap [n]^{(\leqslant r-1)}$ is a non-empty left-compressed up-set in $[n]^{(\leqslant r-1)}$, and so $[r - 1] \in \mathcal{A}'$. Moreover, there is a (U, v, f)-compression with $v = r$ moving $[r - 1]$ to $\{r\} \cup [n - r + 2, n]$. Since all sets added by such a compression contain $v = r$ this does not add $[n - r, n]$, and thus $\mathcal{A}'$ is closed under this compression. Hence $\{r\} \cup [n - r + 2, n] \in \mathcal{A}'$.

Since $\mathcal{A}'$ is left-compressed and $\{r\} \cup [n - r + 2, n] \in \mathcal{A}'$, the family $\mathcal{A}'$ also contains every set that can be obtained from $\{r\} \cup [n - r + 2, n]$ by (repeated) left-compression. This includes every set in $[n]^{(r)}$ containing an element less than or equal to r. It is easy to see that there are strictly more than $\binom{n-1}{r-1}$ such sets and so our claim holds.

Now since $\mathcal{A}'$ contains strictly more than $\binom{n-1}{r-1}$ sets in $[n]^{(r)}$, the Erdős–Ko–Rado theorem tells us that $\mathcal{A}'$ must contain two disjoint sets A and A'. We now complete the proof by applying Lemma 5.2 unless either A or A' contains n.

So assume, without loss of generality, that $n \in A$. Now there must be some $m \neq n$ contained in neither A nor A'. Since $\mathcal{A}'$ is left-compressed, and so in particular mn-compressed, the set

$A'' = A \cup \{m\} \setminus \{n\}$ is also in $\mathcal{A}'$. The set A'' is disjoint from A' and does not contain n. Hence we can apply Lemma 5.2 with sets A'' and A' to conclude the proof in this case.   □

The final few cases

For completeness we prove the final few cases of Theorem 1.2. The arguments in this section are completely different from those given earlier: they are not compression-based. The only remaining cases are

- $n = 2t$ with $2^{n-1} + \frac{1}{2}\binom{n}{t} - t < N < 2^{n-1} + \frac{1}{2}\binom{n}{t}$, and
- $n = 2t + 1$ with $2^{n-1} + \binom{n-1}{t-1} - t - 1 < N < 2^{n-1} + \binom{n-1}{t-1}$.

Any family of size N must contain at least $N - 2^{n-1}$ complementary pairs. For N in the range we are now considering, Katona, Katona and Katona [4] give an example of a family containing exactly this many complementary pairs and no other non-intersecting pairs. Moreover, they observe that any such family contains the maximal number of intersecting subfamilies of size s for every s. Conversely, it is easy to check that if a family $\mathcal{A}$ contains the maximal number of subfamilies of size s for any fixed s, then all non-intersecting pairs in $\mathcal{A}$ must be complementary.

It follows that each set in a complementary pair must be a minimal element of $\mathcal{A}$. Furthermore, the subfamily $\mathcal{B}$ given by

$$\mathcal{B} = \begin{cases} \mathcal{A} \cap [n]^{(\leqslant t)} & n = 2t + 1 \text{ odd} \\ \mathcal{A} \cap ([n]^{(<t)} \cup \{A \in [n]^{(t)} : 1 \in A\}) & n = 2t \text{ even} \end{cases}$$

must be intersecting, since in each case we take the intersection of $\mathcal{A}$ with a family not containing any complementary pairs. Note that $|\mathcal{A}| \leqslant 2^{n-1} + |\mathcal{B}|$ and hence, using the lower bound on N, that $|\mathcal{B}| > \binom{n-1}{t-1} - (n - t)$. Since every set in a complementary pair is minimal and every complementary pair contains an element of $\mathcal{B}$, the following lemma completes the proof.

Lemma 5.3. *Let $\mathcal{B} \subset [n]^{(\leqslant t)}$ be an intersecting family with $\mathcal{B} \not\subset [n]^{(t)}$. Then $\mathcal{B}$ has at most $\binom{n-1}{t-1} - (n - t)$ minimal elements.*

Proof. Let

$$\mathcal{U} = \{B \in \mathcal{B} : B \text{ minimal}, |B| < t\}$$

and

$$\mathcal{V} = \{B \in \mathcal{B} : B \text{ minimal}, |B| = t\}.$$

Let $\partial \mathcal{U} = \{A \in [n]^{(t)} : B \subset A \text{ for some } B \in \mathcal{U}\}$ be the upper shadow of $\mathcal{U}$ in layer t. Note that, by definition of $\mathcal{U}$ and $\mathcal{V}$, the families $\partial \mathcal{U}$ and $\mathcal{V}$ are disjoint. It is easy to see that $\partial \mathcal{U} \cup \mathcal{V}$ is an intersecting family and so, by the Erdős–Ko–Rado theorem, must have size at most $\binom{n-1}{t-1}$. Thus the total number of minimal elements of $\mathcal{B}$ is

$$|\mathcal{V}| + |\mathcal{U}| = |\mathcal{V}| + |\partial \mathcal{U}| + |\mathcal{U}| - |\partial \mathcal{U}| \leqslant \binom{n-1}{t-1} - (|\partial \mathcal{U}| - |\mathcal{U}|),$$

and so it suffices to show that $|\partial \mathcal{U}| - |\mathcal{U}| \geqslant n - t$. As $\mathcal{U}$ is an antichain, it follows easily from the Kruskal–Katona theorem that $\mathcal{U}$ is no larger than its upper shadow in the $(t - 1)$th layer. Hence, we may assume $\mathcal{U} \subset [n]^{(t-1)}$.

Form a bipartite graph G with vertex-sets $[n]^{(t-1)}$ and $[n]^{(t)}$. For $A \in [n]^{(t-1)}$ and $B \in [n]^{(t)}$ we take AB to be an edge of G whenever $A \subset B$. Every vertex in $[n]^{(t-1)}$ has degree $n - t + 1$ and every vertex in $[n]^{(t)}$ has degree t. Fix $A \in \mathcal{U}$ and let $\Gamma(A)$ denote the set of neighbours of A in G. Form a graph G' by deleting $\{A\} \cup \Gamma(A)$ from G. Since every vertex in $[n]^{(t-1)} \setminus \{A\}$ is joined to at most one deleted vertex, we see that the degree in G' of every vertex in $[n]^{(t-1)} \setminus \{A\}$ is at least $n - t$. Hence, since $n - t \geqslant t$, a standard application of Hall's theorem shows that there exists a matching in G' from $[n]^{(t-1)} \setminus \{A\}$ to $[n]^{(t)} \setminus \Gamma(A)$. Thus

$$|\partial \mathcal{U}| \geqslant |\mathcal{U}| - 1 + |\Gamma(A)| = |\mathcal{U}| + n - t. \qquad \square$$

Constructions showing the bound is tight

Recall that a family $\mathcal{A}$ of size N containing precisely $N - 2^{n-1}$ complementary pairs and no other non-intersecting pairs maximizes the number of intersecting subfamilies of every possible size. We remark that an equivalent condition is that $\mathcal{A}$ meets every complementary pair in $\mathcal{P}([n])$ and the only non-intersecting pairs in $\mathcal{A}$ are complementary.

Theorem 5.4. *Suppose that*

(i) $n = 2t$ *and* $2^{n-1} < N \leqslant 2^{n-1} + \frac{1}{2} \binom{n}{t} - t$, *or*

(ii) $n = 2t + 1$ *and* $2^{n-1} < N \leqslant 2^{n-1} + \binom{n-1}{t-1} - t - 1$.

Then there exists a family $\mathcal{A}$ containing the maximal number of intersecting families of every possible size that is not of the form $[n]^{(\geqslant t+1)} \cup \mathcal{B}$ for any $\mathcal{B} \subset [n]^{(t)}$.

Proof. (i) Consider the family

$$\mathcal{A} = \{[t-1]\} \cup [n]^{(\geqslant t)} \setminus \{A \in [n]^{(t)} : A \cap [t-1] = \emptyset\}.$$

This family has size $2^{n-1} + \frac{1}{2} \binom{n}{t} - t$ and meets every complementary pair in $\mathcal{P}([n])$, and the only non-intersecting pairs in $\mathcal{A}$ are complementary pairs. Thus, $\mathcal{A}$ satisfies the conclusion of the theorem for $N = 2^{n-1} + \frac{1}{2} \binom{n}{t} - t$. Moreover, for $2^{n-1} < N < 2^{n-1} + \frac{1}{2} \binom{n}{t} - t$ we may obtain a suitable family by deleting from $\mathcal{A}$ the appropriate number of sets in $[n]^{(t)}$ that contain the element 1. (ii) In this case we apply an identical argument, starting from the family

$$\mathcal{A} = \{[t-1]\} \cup \{A \in [n]^{(t)} : 1 \in A\} \cup [n]^{(\geqslant t+1)} \setminus \{A \in [n]^{(t+1)} : A \cap [t-1] = \emptyset\}$$

of size $2^{n-1} + \binom{n-1}{t-1} - t - 1$ and deleting sets in $[n]^{(t)}$ that contain the element 1 but do not contain all of $[t-1]$. $\qquad \square$

6. Concluding remarks and open questions

As we remarked earlier, there seems to be no reason to believe that the maximizing families for different values of p are the same. However, in all cases where the maximizing families are known, including the new examples in this paper, they are in fact the same for all p. Even more is true: the known examples simultaneously maximize the number of intersecting subfamilies of every possible size. We therefore recall the following question, first asked in [5].

Question. Suppose $N > 2^{n-1}$. Does there exist a family $\mathcal{A}$ of size N which simultaneously maximizes the number of intersecting subfamilies of size s for every s?

One could, of course, ask an analogous question for families $\mathcal{A}$ restricted to lie in a single layer of the cube $\mathcal{P}([n])$. But here it is not always possible to simultaneously maximize the number of intersecting families of every size. Indeed, we have seen that in $[n]^{(2)}$ the family of size $\binom{n-1}{2}$ with the most intersecting pairs is $\{A \in [n]^{(2)} : n \notin A\}$. However, this is obviously not the family containing the most intersecting subfamilies of size n: it does not contain any at all.

The exact extremal families for most values of N remain unknown. Indeed, even the following question is open.

Question. For every N satisfying the lower bound of Theorem 1.2, is there a unique (up to reordering the coordinates) family $\mathcal{A}$ that maximizes the probability that $\mathcal{A}_p$ is intersecting?

The family is not unique for N less than the bound given: for example, the optimal family constructed in Theorem 5.4 is not the same as that constructed in [4]. Theorem 1.2 shows that the family is unique for certain values of N, primarily N of the form $N = \sum_{k=r}^{n} \binom{n}{k}$.

We remark that, for fixed s, the family containing the most intersecting subfamilies of size s is not always unique: indeed, for $s = 2$, $n \equiv 0, 1 \bmod 4$ and $N = \sum_{k=3}^{n} \binom{n}{k} + \frac{1}{2}\binom{n}{2}$, the results of Ahlswede and Katona [2] imply that the families obtained by taking the union of $[n]^{(\geqslant 3)}$ with a quasi-clique or a quasi-star are both optimal. However, it is easy to check that, at least for large n, the family with the quasi-star has far more intersecting triples and, indeed, far more intersecting families of size s for all $s > 2$. Hence the family maximizing the probability that $\mathcal{A}_p$ is intersecting is unique in this case.

References

[1] Ahlswede, R. (1980) Simple hypergraphs with maximal number of adjacent pairs of edges. *J. Combin. Theory Ser. B* **28** 164–167.
[2] Ahlswede, R. and Katona, G. O. H. (1978) Graphs with maximal number of adjacent pairs of edges. *Acta Math. Acad. Sci. Hungar.* **32** 97–120.
[3] Frankl, P. (1977) On the minimum number of disjoint pairs in a family of finite sets. *J. Combin. Theory Ser. A* **22** 249–251.
[4] Katona, G. O. H., Katona, G. Y. and Katona, Z. (2012) Most probably intersecting families of subsets. *Combin. Probab. Comput.* **21** 219–227.
[5] Russell, P. A. (2012) Compressions and probably intersecting families. *Combin. Probab. Comput.* **21** 301–313.
[6] Wagner, S. and Wang, H. (2009) On a problem of Ahlswede and Katona. *Studia Sci. Math. Hungar.* **46** 423–435.