G-intersection theorems for matchings and other graphs

J Robert Johnson[*] John Talbot[†]

July 30, 2009

Abstract

If G is a graph with vertex set $[n]$ then $\mathcal{A} \subseteq 2^{[n]}$ is G-*intersecting* if for all $A, B \in \mathcal{A}$ either $A \cap B \neq \emptyset$ or there exist $a \in A$ and $b \in B$ such that $a \sim_G b$.

The question of how large a k-uniform G-intersecting family can be was first considered by Bohman, Frieze, Ruszinkó and Thoma [2] who identified two natural candidates for the extrema depending on the relative sizes of k and n and asked whether there is a sharp phase transition between the two. Our first result shows that there is a sharp transition and characterizes the extremal families when G is a matching. We also give an example demonstrating that other extremal families can occur.

Our second result gives a sufficient condition for the largest G-intersecting family to contain almost all k-sets. In particular we show that if C_n is the n-cycle and $k > \alpha n + o(n)$, where $\alpha = 0.266\ldots$ is the smallest positive root of $(1-x)^3(1+x) = 1/2$, then the largest C_n-intersecting family has size $(1-o(1))\binom{n}{k}$.

Finally we consider the non-uniform problem and show that in this case the size of the largest G-intersecting family depends on the matching number of G.

1 Introduction

The following generalization of the notion of an intersecting family was introduced by Bohman, Frieze, Ruszinkó and Thoma [2]. If G is a graph with vertex set $[n]$ then $\mathcal{A} \subseteq 2^{[n]}$ is G-*intersecting* if for all $A, B \in \mathcal{A}$ either $A \cap B \neq \emptyset$ or there exist $a \in A$ and $b \in B$ such that $a \sim_G b$.

The question of how large a k-uniform G-intersecting family can be is a natural generalization of the Erdős–Ko–Rado problem, indeed if G is the empty graph it is answered by the classical Erdős–Ko–Rado theorem [5].

[*]School of Mathematical Sciences, Queen Mary, University of London, E1 4NS, UK, Email: r.johnson@qmul.ac.uk

[†]Department of Mathematics, University College London, WC1E 6BT, UK, Email: talbot@math.ucl.ac.uk. This author is a Royal Society University Research Fellow.

Theorem 1.1 (Erdős–Ko–Rado 1938 [5])**.** *If $\mathcal{A} \subseteq \binom{[n]}{k}$ is intersecting then*

$$|\mathcal{A}| \leq \begin{cases} \binom{n-1}{k-1}, & 1 \leq k \leq n/2, \\ \binom{n}{k}, & n/2 < k \leq n. \end{cases}$$

Moreover if $k < n/2$ then equality is attained iff $\mathcal{A}$ consists of all k-sets containing a fixed element of $[n]$. While (trivially) if $k > n/2$ then equality is attained iff $\mathcal{A} = \binom{[n]}{k}$.

For a graph G with vertex set $[n]$ and $1 \leq k \leq n$ we define

$$N(G,k) = \max\left\{|\mathcal{A}| : \mathcal{A} \subseteq \binom{[n]}{k} \text{ is } G\text{-intersecting}\right\}.$$

Bohman et al. [2] were the first to consider the problem of determining $N(G,k)$. They identified two types of behaviour for the extrema depending on the relative sizes of k and n, mirroring the extremal behaviour of ordinary k-uniform intersecting families (as given by Theorem 1.1).

The *augmented neighbourhood* of $A \subseteq [n]$, denoted by $\Gamma^+(A)$, is the union of A and its neighbourhood in G. So a family $\mathcal{A} \subseteq 2^{[n]}$ is G-intersecting iff for all $A, B \in \mathcal{A}$ we have $A \cap \Gamma^+(B) \neq \emptyset$.

An obvious example of a k-uniform G-intersecting family is the collection of all k-sets meeting a fixed clique in G. For instance if $G = C_n$ is the n-cycle then

$$\mathcal{A} = \left\{A \in \binom{[n]}{k} : A \cap \{1,2\} \neq \emptyset\right\},$$

is C_n-intersecting. However $\mathcal{A}$ is not maximal: it can be extended to

$$\mathcal{B} = \mathcal{A} \cup \left\{B \in \binom{[n]}{k} : 3, n \in B\right\}.$$

More generally if K is a clique in G and $M_1, M_2, \ldots, M_r \subseteq [n]\backslash K$ satisfy

$$K \subseteq \Gamma^+(M_i) \text{ for } 1 \leq i \leq r \quad \text{and} \quad M_i \cap \Gamma^+(M_j) \neq \emptyset,\ i \neq j,$$

then

$$\mathcal{A}(K; M_1, \ldots, M_r) = \left\{A \in \binom{[n]}{k} : A \cap K \neq \emptyset \text{ or } M_i \subseteq A \text{ for some } i\right\}, \quad (1)$$

is also G-intersecting. We will call such a family a (G,k)-*star* with *centre*

$$C = K \cup \bigcup_{i=1}^{r} M_i.$$

Bohman et al. [2] showed that if G is sparse and $k = O(n^{1/4})$ then the largest G-intersecting families are of this form. (More recently Bohman and Martin [3] gave an improvement, showing that a similar result also holds for $k = O(n^{1/2})$.)

Bohman et al. [2] also showed that if G is sparse with minimum degree δ and $k > cn$, where c is a constant satisfying $c - (1-c)^{\delta+1} > 0$, then

$$N(G, k) = (1 - o(1))\binom{n}{k}.$$

These two different types of extrema mirror the two cases of the Erdős–Ko–Rado theorem, however there is a large gap between the values of k for which they are known to occur. Bohman et al. [2] asked whether there is a sharp phase transition and whether other types of extrema exist.

Our first result in the next section (Theorem 2.1) shows that there is a sharp transition and characterizes the extremal families when G is a perfect matching. We also give an example of a graph demonstrating that other types of extrema exist.

In the third section we give a sufficient condition for the largest G-intersecting family to contain almost all k-sets (Theorem 3.1). In particular we show that if C_n is the n-cycle and $k > \alpha n + o(n)$, where $\alpha = 0.266\ldots$ is the smallest positive root of $(1-x)^3(1+x) = 1/2$, then the largest C_n-intersecting family has size $(1-o(1))\binom{n}{k}$ (Corollary 3.2). This improves an earlier bound of $k > 0.317n$ due to Bohman et al. [2].

In the fourth section we consider the non-uniform problem and show that in this case the size of the largest G-intersecting family depends on the matching number of G (Theorem 4.1).

We end the paper with some open problems and conjectures.

2 Matchings

Let M_n be a matching of order $n = 2t$ with edges $e_1, \ldots, e_t$, where $e_i = \{2i-1, 2i\}$. For $A \in \binom{[n]}{k}$ let $I_A = \{i \in [t] : A \cap e_i \neq \emptyset\}$ (so I_A indexes the edges that A meets). An obvious candidate for the largest M_n-intersecting family when k is small is

$$\mathcal{A}_{\text{pair}} = \left\{ A \in \binom{[n]}{k} : 1 \in I_A \right\}.$$

The precise form of the extremal family when k is large will depend on the parity of t. For t odd let

$$\mathcal{A}_{\text{maj}} = \left\{ A \in \binom{[n]}{k} : |I_A| > \frac{t}{2} \right\}.$$

For t even we can extend $\mathcal{A}_{\text{maj}}$ by adding half of those k-sets meeting exactly $t/2$ edges. To be precise, for t even let $\mathcal{B} \subseteq \binom{t}{t/2}$ be an (ordinary) intersecting family of maximum size $\frac{1}{2}\binom{t}{t/2}$. We define

$$\mathcal{A}_{\text{maj}} = \left\{ A \in \binom{[n]}{k} : |I_A| > \frac{t}{2} \text{ or } I_A \in \mathcal{B} \right\}.$$

Note that both $\mathcal{A}_{\text{pair}}$ and $\mathcal{A}_{\text{maj}}$ are M_n-intersecting.

The result of Bohman and Martin (Theorem 2 [3]) implies that $\mathcal{A}_{\text{pair}}$ is a k-uniform M_n-intersecting family of maximum size for $k = O(n^{1/2})$ while the result of Bohman et al. (Theorem 7 [2]) implies that $N(M_n, k) = (1 - o(1))\binom{n}{k}$ for $k > 0.38196n$. We are able to give the following result describing a sharp threshold for the behaviour of $N(M_n, k)$ and characterizing the extremal families.

Theorem 2.1. *Let $n = 2t \geq 1000$, $1 \leq k \leq n$ and M_n be a matching of order n with edges $\{1, 2\}$, $\{3, 4\}, \ldots, \{n-1, n\}$. If $d = 1 - 2^{-1/2} = 0.29289\ldots$ then*

$$N(M_n, k) = \begin{cases} |\mathcal{A}_{pair}| = \binom{n}{k} - \binom{n-2}{k}, & k < dn, \\ |\mathcal{A}_{maj}| = (1 - o(1))\binom{n}{k}, & k > dn(1 + \epsilon_n), \end{cases}$$

where $\epsilon_n = 30\sqrt{\frac{\log n}{n}} = o(1)$. Moreover, up to isomorphism, these bounds are only achieved by the families $\mathcal{A}_{pair}$ and $\mathcal{A}_{maj}$ described above.

For the remainder of this section we will say that k is *small* (with respect to n) if $k < dn$ and k is *large* (with respect to n) if $k > dn(1 + \epsilon_n)$.

Proof of Theorem 2.1. Let $\mathcal{A} \subseteq \binom{[n]}{k}$ be a k-uniform M_n-intersecting family of maximum size (so $|\mathcal{A}| = N(M_n, k)$). We define

$$\mathcal{I}(\mathcal{A}) = \{I_A \subseteq [t] : A \in \mathcal{A}\}.$$

Note that the sets in $\mathcal{I}(\mathcal{A})$ all have sizes in the range $\lceil k/2 \rceil$ up to k. (Since a k-set cannot meet less than $\lceil k/2 \rceil$ edges or more than k edges.)

For $B \subseteq [t]$ define

$$W_k(B) = \left\{ A \in \binom{[n]}{k} : I_A = B \right\}.$$

So $W_k(B)$ is the family of all k-sets meeting precisely those edges indexed by B. The size of this family depends only on the size of B. For $1 \leq m \leq t$ let $w_k(m) = |W_k([m])|$. (So $w_k(m) \neq 0$ iff $\lceil k/2 \rceil \leq m \leq k$.)

We note a few useful facts whose proofs we defer.

Lemma 2.2. *$\mathcal{I}(\mathcal{A})$ has the following properties:*

(a) $\mathcal{I}(\mathcal{A})$ is intersecting.

(b) If $B \subseteq [t]$, $\lceil k/2 \rceil \leq |B| \leq k$ and $\mathcal{I}(\mathcal{A}) \cup \{B\}$ is intersecting then $B \in \mathcal{I}(\mathcal{A})$.

(c) If $B \in \mathcal{I}(\mathcal{A})$, $B \subset C \subseteq [t]$ and $|C| \leq k$ then $C \in \mathcal{I}(\mathcal{A})$.

Lemma 2.3. *If $A \in \mathcal{A}$, $B \in \binom{[n]}{k}$ and $I_A = I_B$ then $B \in \mathcal{A}$. Hence*

$$\mathcal{A} = \dot{\bigcup_{B \in \mathcal{I}(\mathcal{A})}} W_k(B).$$

Thus if $\mathcal{A}_m = \{A \in \mathcal{A} : |I_A| = m\}$, $\mathcal{I}_m(\mathcal{A}) = \{B \in \mathcal{I}(\mathcal{A}) : |B| = m\}$ *and* $i_m(\mathcal{A}) = |\mathcal{I}_m(\mathcal{A})|$ *then*

$$|\mathcal{A}_m| = i_m(\mathcal{A})w_k(m) \qquad \textit{and} \qquad |\mathcal{A}| = \sum_{m=\lceil k/2\rceil}^{k} i_m(\mathcal{A})w_k(m).$$

Using Lemmas 2.2 and 2.3 the problem of determining $N(M_n, k)$ can be reduced to a weighted intersection problem for $\mathcal{I}(\mathcal{A})$, with the weight of a set $B \in \mathcal{I}(\mathcal{A})$ given by $w_k(|B|)$.

Lemma 2.4.

$$w_k(m) = \begin{cases} 2^{2m-k}\binom{m}{k-m}, & \lceil k/2\rceil \le m \le k, \\ 0, & \textit{otherwise.} \end{cases}$$

Lemma 2.5. *There exists* $m^* = m^*(k,t)$ *such that* $\{w_k(m)\}_{m=\lceil k/2\rceil}^{k}$ *satisfies*

$$w_k(\lceil k/2\rceil) < \cdots < w_k(m^*) \ge w_k(m^*+1) > w_k(m^*+2) > \cdots > w_k(k).$$

That is the sequence is strictly increasing up to a maximum which is attained at m^* *and possibly also attained at* m^*+1 *and thereafter the sequence is strictly decreasing.*

Lemma 2.6. *If* k *is small then* $m^* < t/2$. *Moreover if* $\lceil k/2\rceil \le m \le k$ *and* $m < t/2$ *then* $w_k(m) > w_k(t-m)$.

Lemma 2.7. *If* k *is large then* $m^* > t/2$ *and*

$$\sum_{m<t/2} w_k(m)\binom{t}{m} < \frac{1}{t}\binom{n}{k}.$$

Lemma 2.8. *If* $m_1 \le m_2$, $\mathcal{B} \subseteq \binom{[t]}{m_2}$ *and*

$$\partial^{(m_1)}(\mathcal{B}) = \left\{C \in \binom{[t]}{m_1} : C \subseteq B \textit{ for some } B \in \mathcal{B}\right\}$$

then

$$\frac{|\partial^{(m_1)}(\mathcal{B})|}{\binom{t}{m_1}} \ge \frac{|\mathcal{B}|}{\binom{t}{m_2}}.$$

Returning to the proof of Theorem 2.1 we suppose first that k is small and let $\mathcal{I}(\mathcal{A})$, $\mathcal{A}_m$, $\mathcal{I}_m(\mathcal{A})$ and $i_m(\mathcal{A})$ be as defined above. Lemma 2.3 implies that for $m < t/2$

$$|\mathcal{A}_m| + |\mathcal{A}_{t-m}| = i_m(\mathcal{A})w_k(m) + i_{t-m}(\mathcal{A})w_k(t-m). \tag{2}$$

By Lemma 2.2 (a), $i_m(\mathcal{A}) + i_{t-m}(\mathcal{A})$ is the size of an intersecting family in $\binom{[t]}{m} \cup \binom{[t]}{t-m}$. Hence Lemma 2.6 and the Erdős–Ko–Rado theorem (Theorem 1.1) imply that

$$|\mathcal{A}_m| + |\mathcal{A}_{t-m}| \le \binom{t-1}{m-1}w_k(m) + \binom{t-1}{t-m-1}w_k(t-m),$$

with strict inequality unless $\mathcal{I}_m(\mathcal{A}) \cup \mathcal{I}_{t-m}(\mathcal{A})$ consists of all sets in $\binom{[t]}{m} \cup \binom{[t]}{t-m}$ containing a fixed element of $[t]$ (which is the case for $\mathcal{A}_{\text{pair}}$). Finally if t is even then Theorem 1.1 implies that

$$|\mathcal{A}_{t/2}| \leq w_k(t/2) \frac{1}{2} \binom{t}{t/2} \tag{3}$$

and again this is achieved by $\mathcal{A}_{\text{pair}}$. Hence $N(M_n, k) = |\mathcal{A}| \leq |\mathcal{A}_{\text{pair}}|$.

To see that $\mathcal{A}_{\text{pair}}$ is (up to isomorphism) the unique extremal family note that for equality to hold $\mathcal{I}_{\lceil k/2 \rceil}(\mathcal{A})$ must consist of all $\lceil k/2 \rceil$-sets containing a fixed element $i \in [t]$. Without loss of generality we may suppose that $i = 1$. Lemma 2.2 (c) now implies that $\mathcal{A}_{\text{pair}} \subseteq \mathcal{A}$. Finally $k < dn$ implies that $\mathcal{A}_{\text{pair}}$ is a maximal M_n-intersecting family and hence $\mathcal{A} = \mathcal{A}_{\text{pair}}$. (Indeed it is easy to check that $\mathcal{A}_{\text{pair}}$ is a maximal M_n-intersecting family whenever $k \leq n/3$.) Note that this part of Theorem 2.1 holds for all values of n and $k < dn$ (the condition $n \geq 1000$ is only required for the k large case).

Now suppose that k is large, that is $k > dn(1 + \epsilon_n)$. If $k > t = n/2$ then $\mathcal{A}_{\text{maj}} = \binom{[n]}{k}$ and the result is trivial, so suppose that $k \leq t$. In this case Lemma 2.7 tells us that the maximum of $w_k(m)$ is achieved at $m^* > t/2$. By Lemma 2.5, $w_k(m)$ is strictly increasing for $m < m^*$ so $w_k(m) > w_k(t - m)$ for $t/2 < m \leq m^*$. By Lemma 2.2 (a), $i_m(\mathcal{A}) + i_{t-m}(\mathcal{A})$ is the size of an intersecting family in $\binom{[t]}{m} \cup \binom{[t]}{t-m}$ so, for $t/2 < m \leq m^*$, $w_k(m) > w_k(t - m)$ implies that

$$|\mathcal{A}_m| + |\mathcal{A}_{t-m}| \leq \binom{t}{m} w_k(m), \tag{4}$$

with strict inequality unless $\mathcal{I}_{t-m}(\mathcal{A})$ is empty and $\mathcal{I}_m(\mathcal{A}) = \binom{[t]}{m}$ (which is the case for $\mathcal{A}_{\text{maj}}$).

If t is odd then (4) bounds the number of sets in $\mathcal{A}_m$ for $t - m^* \leq m \leq m^*$. However if t is even then we note that $i_{t/2}(\mathcal{A})$ is the size of an intersecting family in $\binom{t}{t/2}$ and so (3) holds.

If $\mathcal{I}_m(\mathcal{A}) = \emptyset$ for $m < t - m^*$ then (4) and, in the case of t even (3), imply that $|\mathcal{A}| \leq |\mathcal{A}_{\text{maj}}|$, with strict inequality unless $\mathcal{A} = \mathcal{A}_{\text{maj}}$. We suppose now, for a contradiction, that there exists $m < t - m^*$ such that $\mathcal{I}_m(\mathcal{A}) \neq \emptyset$. Let $m_0 < t - m^*$ be chosen so that

$$\frac{i_{m_0}(\mathcal{A})}{\binom{t}{m_0}} = \max \left\{ \frac{i_m(\mathcal{A})}{\binom{t}{m}} : m < t - m^* \right\}.$$

(So the proportion of m_0-sets in $\mathcal{I}(\mathcal{A})$ is maximal subject to $m_0 < t - m^*$.) Let $\beta = i_{m_0}(\mathcal{A}) / \binom{t}{m_0} > 0$. By Lemma 2.3 we have

$$\sum_{m < t - m^*} |\mathcal{A}_m| \leq \beta \sum_{m < t - m^*} w_k(m) \binom{t}{m}. \tag{5}$$

The complements of the sets in $\mathcal{I}_{m_0}(\mathcal{A})$ are all missing from $\mathcal{I}_{t-m_0}(\mathcal{A})$, so if $\mathcal{B} = \binom{[t]}{t-m_0} \setminus \mathcal{I}_{t-m_0}(\mathcal{A})$ then $|\mathcal{B}| \geq i_{m_0}(\mathcal{A}) = \beta\binom{t}{m_0}$. Recall that

$$\partial^{(m)}(\mathcal{B}) = \left\{ C \in \binom{[t]}{m} : C \subseteq B \text{ for some } B \in \mathcal{B} \right\}.$$

If $t/2 < m \leq t - m_0$ then Lemma 2.2 (c) implies that $\partial^{(m)}(\mathcal{B}) \cap \mathcal{I}_m(\mathcal{A}) = \emptyset$ while Lemma 2.8 implies that

$$\frac{|\partial^{(m)}(\mathcal{B})|}{\binom{t}{m}} \geq \frac{|\mathcal{B}|}{\binom{t}{t-m_0}} \geq \beta.$$

Note that $m^* < t - m_0$ and hence for $t/2 < m \leq m^*$,

$$i_m(\mathcal{A}) \leq \binom{t}{m} - |\partial^{(m)}(\mathcal{B})| \leq (1-\beta)\binom{t}{m}. \tag{6}$$

Now if $t/2 < m \leq m^*$ then $w_k(m) > w_k(t-m)$, so (6) implies that

$$|\mathcal{A}_m| + |\mathcal{A}_{t-m}| \leq \binom{t}{m}((1-\beta)w_k(m) + \beta w_k(t-m)). \tag{7}$$

If $m < t - m^*$ then, by the definition of β,

$$|\mathcal{A}_m| \leq \beta\binom{t}{m} w_k(m). \tag{8}$$

While trivially if $m > m^*$ then $|\mathcal{A}_m| \leq w_k(m)\binom{t}{m}$. Together with (7), (8), and in the case of t even (3), this implies that

$$|\mathcal{A}| - |\mathcal{A}_{\text{maj}}| \leq \sum_{m<t/2} \beta\binom{t}{m} w_k(m) - \sum_{t/2<m\leq m^*} \beta\binom{t}{m} w_k(m).$$

Now Lemma 2.7 implies that

$$|\mathcal{A}| - |\mathcal{A}_{\text{maj}}| < \beta \left(\frac{1}{t}\binom{n}{k} - \sum_{t/2<m\leq m^*} \binom{t}{m} w_k(m) \right).$$

So we will have$|\mathcal{A}| < |\mathcal{A}_{\text{maj}}|$ (and the proof will be complete) if we show that

$$\sum_{t/2<m\leq m^*} \binom{t}{m} w_k(m) \geq \frac{1}{t}\binom{n}{k}. \tag{9}$$

Let m_1 satisfy

$$w_k(m_1)\binom{t}{m_1} = \max\left\{ w_k(m)\binom{t}{m} : \lceil k/2 \rceil \leq m \leq k \right\}.$$

Recall that
$$\sum_{m=\lceil k/2\rceil}^{k}\binom{t}{m}w_k(m)=\binom{n}{k}$$
and $k\le t$. Hence
$$\binom{t}{m_1}w_k(m_1)\ge\frac{1}{\frac{k}{2}+1}\binom{n}{k}\ge\frac{2}{t+2}\binom{n}{k}. \tag{10}$$
Since $m^*>t/2$, both $w_k(m)$ and $\binom{t}{m}$ are increasing in m for $m\le t/2$ and decreasing in m for $m>m^*$ so $m_1\in\{\lceil t/2\rceil,\dots,m^*\}$. If $m_1\ne t/2$ then $t/2<m_1\le m^*$ and so (10) implies that (9) holds. Otherwise t is even and $m_1=t/2$. In which case (10) and $w_k(t/2+1)>w_k(t/2)$ imply that
$$\binom{t}{t/2+1}w_k(t/2+1)>\frac{t}{t+2}\binom{t}{t/2}w_k(t/2)\ge\frac{2t}{(t+2)^2}\binom{n}{k}.$$
This implies that (9) holds so long as $2t^2\ge(t+2)^2$, which is true for $t\ge5$. $\square$

Proof of Lemma 2.2. If $A,B\in\mathcal{I}(\mathcal{A})$ then there exist $C,D\in\mathcal{A}$ such that $A=I_C$ and $B=I_D$. Since $\mathcal{A}$ is M_n-intersecting so C and D must meet a common edge e_i. Hence $i\in A\cap B$ and so $\mathcal{I}(\mathcal{A})$ is intersecting.

If Lemma 2.2 (b) does not hold then there is $B\subseteq[t]$, with $\lceil k/2\rceil\le|B|\le k$ such that $\mathcal{I}(\mathcal{A})\cup\{B\}$ is intersecting but $B\notin\mathcal{I}(\mathcal{A})$. Now $\lceil k/2\rceil\le|B|\le k$ implies that $W_k(B)\ne\emptyset$ and so $\mathcal{B}=\mathcal{A}\cup W_k(B)$ is a k-uniform M_n-intersecting family satisfying $|\mathcal{B}|>|\mathcal{A}|$, contradicting the maximality of $|\mathcal{A}|$.

Finally (b) implies (c) since if $B\subset C$ and $B\in\mathcal{I}(\mathcal{A})$ then $\mathcal{I}(\mathcal{A})\cup\{C\}$ is also intersecting. $\square$

Proof of Lemma 2.3. If $A\in\mathcal{A}$, $B\in\binom{[n]}{k}$ and $I_A=I_B$ then $\mathcal{A}\cup\{B\}$ is M_n-intersecting. So, by the maximality of $|\mathcal{A}|$, $B\in\mathcal{A}$. Hence
$$\mathcal{A}=\bigcup_{B\in\mathcal{I}(\mathcal{A})}W_k(B).$$
Moreover this is a disjoint union since if $B,C\in\mathcal{I}(\mathcal{A})$ and $B\ne C$ then $W_k(B)\cap W_k(C)=\emptyset$. The final part follows directly from the definitions. $\square$

Proof of Lemma 2.4. If $m<\lceil k/2\rceil$ or $m>k$ then no k-set meets m edges and hence $w_k(m)=0$, so suppose that $\lceil k/2\rceil\le m\le k$. Consider $A\in\binom{[n]}{k}$ meeting the first m edges of M_n (that is $A\in W_k([m])$). For such a set let a_i denote the number of edges it meets in exactly i elements, where $i=1,2$. Since $a_1+a_2=m$ and $a_1+2a_2=k$ we have $a_1=2m-k$ and $a_2=k-m$. Thus such a set is uniquely determined by choosing $k-m$ of the m edges from which to take both vertices and then choosing one of the two possible vertices from each of the remaining $2m-k$ edges. $\square$

Proof of Lemma 2.5. We show first that if $\lceil k/2 \rceil \leq m \leq k$ then

$$w_k(m)^2 > w_k(m+1)w_k(m-1). \tag{11}$$

First note that this holds for $m = \lceil k/2 \rceil$ or $m = k$ since in this case the RHS of (11) is zero while the LHS is positive. So suppose that $\lceil k/2 \rceil < m < k$. Now

$$\begin{aligned}
\frac{w_k(m)^2}{w_k(m+1)w_k(m-1)} &= \frac{m(k-m+1)(2m-k+2)(2m-k+1)}{(m+1)(k-m)(2m-k)(2m-k-1)} \\
&> \frac{m(k-m+1)}{(m+1)(k-m)} \\
&> 1.
\end{aligned}$$

Hence if $y_m = w_k(m)/w_k(m+1)$ then $\{y_m\}_{m=\lceil k/2 \rceil}^{k}$ is strictly increasing. This implies the result. $\square$

Proof of Lemma 2.6. Let k be small. We will assume that $t = 2s$ is even, the proof for t odd is essentially identical. By Lemma 2.5 if $w_k(s-1) > w_k(s)$ then $m^* < s$. Recall that since k is small we have $k \leq dn - 1$, where $d = 1 - 2^{-1/2}$ and $n = 2t = 4s$. Now

$$\begin{aligned}
\frac{w_k(s-1)}{w_k(s)} &= \frac{(2s-k)(2s-k-1)}{4s(k-s+1)} \\
&> \frac{(2s-k-1)^2}{4s(k-s+1)} \\
&\geq 1.
\end{aligned}$$

Hence $m^* < s = t/2$.

We now prove by induction on $a \geq 1$ that if $\lceil k/2 \rceil \leq s - a \leq k$ then $w_k(s-a) > w_k(s+a)$. (Note that we may suppose that $s + a \leq k$ since otherwise $w_k(s+a) = 0$.) For $a = 1$ this follows from $m^* \leq s - 1$ and Lemma 2.5 so suppose that $a \geq 2$ and the result holds for $a - 1$. It is sufficient to show that

$$\frac{w_k(s-a)}{w_k(s+a)} \geq \frac{w_k(s-(a-1))}{w_k(s+a-1)}, \tag{12}$$

since the RHS of (12) is strictly greater than 1 by our inductive hypothesis. We consider

$$\gamma = \frac{w_k(s-a)w_k(s+a-1)}{w_k(s+a)w_k(s-(a-1))}.$$

We wish to show that $\gamma \geq 1$. Now

$$\begin{aligned}
\gamma &= \frac{(2s-2a-k+2)(2s-2a-k+1)(2s+2a-k)(2s+2a-k-1)}{16(s-a+1)(s+a)(k-s+a)(k-s-a+1)} \\
&> \frac{(2s-k+2a-1)^2(2s-k-2a-1)^2}{16((s^2-a^2)((k-s+1)^2-a^2)} \qquad (13)
\end{aligned}$$

Since (13) is decreasing in k and $k < dn$ (since k is small) we may suppose that $k = dn - 1 = 4ds - 1$. Rearranging we now need to check that

$$(s-2ds+a)^2(s-2ds-a)^2 - (s^2-a^2)((4ds-s)^2-a^2) \geq 0. \qquad (14)$$

Differentiating (14) with respect to a we see that it is increasing in a, for $a > 0$ (the partial derivative is $16ad^2s^2$). Hence it is sufficient to check that (14) holds for $a = 0$, which it does with equality. □

For the proof of Lemma 2.7 we will require Azuma's inequality.

Lemma 2.9 (Azuma [1]). *If $Y_0, \ldots, Y_t$ is a martingale and $|Y_i - Y_{i-1}| \leq c_i$ for $1 \leq i \leq t$ then for any $\lambda > 0$*

$$\mathbb{P}(Y_t \geq Y_0 + \lambda) \leq \exp\left(\frac{-\lambda^2}{2\sum_{i=1}^t c_i^2}\right).$$

Proof of Lemma 2.7. Let k be large. We will assume that $t = 2s$ is even, the proof for t odd is essentially identical. By Lemma 2.5 if $w_k(s) < w_k(s+1)$ then $m^* > s$. Since k is large we have $k \geq dn + 4d = 4d(s+1)$, where $d = 1 - 2^{-1/2}$ and $n = 2t = 4s$. Now

$$\begin{aligned} \frac{w_k(s)}{w_k(s+1)} &= \frac{(2s-k+2)(2s-k+1)}{4(s+1)(k-s)} \\ &< \frac{(2s-k+2)^2}{4(s+1)(k-s-1)} \\ &\leq 1. \end{aligned}$$

Hence $m^* > s = t/2$.

We now need to show that

$$\sum_{m<t/2} w_k(m)\binom{t}{m} < \frac{1}{t}\binom{n}{k}. \qquad (15)$$

If $A \in \binom{[n]}{k}$ is chosen uniformly at random and

$$I_A = \{i \in [t] : A \cap e_i \neq \emptyset\},$$

then

$$\mathbb{P}(|I_A| < t/2) = \sum_{m<t/2} \frac{w_k(m)\binom{t}{m}}{\binom{n}{k}}.$$

So it is sufficient to prove that

$$\mathbb{P}(|I_A| < t/2) < \frac{1}{t}. \qquad (16)$$

For $j \in [n]$ let

$$X_j(A) = \begin{cases} 1, & j \in A, \\ 0, & \text{otherwise.} \end{cases}$$

Recall that M_n has edges $e_1, \dots, e_t$, with $e_i = \{2i-1, 2i\}$. For $i \in [t]$ let $\mathcal{F}_i$ be the σ-algebra generated by $X_1, X_2, \dots, X_{2i}$ and define $Y_i = \mathbb{E}(|[t]\backslash I_A| \mid \mathcal{F}_i)$. If $Y_0 = \mathbb{E}(|[t]\backslash I_A|)$ then $Y_0, Y_1, \dots, Y_t$ is a martingale and

$$\mathbb{E}(|[t]\backslash I_A|) = \sum_{i=1}^{t} \mathbb{P}(A \cap e_i = \emptyset) = \frac{t\binom{n-2}{k}}{\binom{n}{k}}.$$

The values of X_{2i-1} and X_{2i} can change the expected number of edges which A meets among $e_{i+1}, \dots, e_t$ by at most two, as well as determining whether or not $i \in I_A$. Hence

$$|Y_i - Y_{i-1}| \leq 3.$$

Azuma's inequality then implies that

$$\mathbb{P}(Y_t > t\frac{\binom{n-2}{k}}{\binom{n}{k}} + \sqrt{18t\log t}) < \frac{1}{t}.$$

Now $|I_A| = t - Y_t$ so the proof will be complete if we show that (for k large)

$$\frac{t\binom{n-2}{k}}{\binom{n}{k}} + \sqrt{18t\log t} \leq \frac{t}{2}.$$

This will hold if

$$(2t-k)^2 \leq 2t^2 - 4t^2\sqrt{\frac{18\log t}{t}}.$$

A routine calculation now shows that this holds for $k > dn\left(1 + 30\sqrt{\frac{\log n}{n}}\right)$ and $n = 2t \geq 1000$. □

Proof of Lemma 2.8. This is a simple exercise in double counting. Each set $B \in \mathcal{B}$ contains $\binom{m_2}{m_1}$ subsets of size m_1, while each set $C \in \partial^{(m_1)}(\mathcal{B})$ is contained in $\binom{t-m_1}{m_2-m_1}$ supersets of size m_2 (and thus in at most this number of sets in $\mathcal{B}$). Hence

$$|\partial^{(m_1)}(\mathcal{B})|\binom{t-m_1}{m_2-m_1} \geq |\mathcal{B}|\binom{m_2}{m_1},$$

which implies the result. □

In fact the same value $d = 1 - 2^{-1/2}$ is a threshold for a slightly more general class of graphs.

Theorem 2.10. *If G_n is the graph of order n with w_n pairwise disjoint edges and $n - 2w_n$ isolated vertices, where $w_n \to \infty$ as $n \to \infty$, then there exists $\delta_n = o(1)$ such that*

$$N(G_n, k) = \begin{cases} \binom{n}{k} - \binom{n-2}{k}, & k < dn, \\ (1-o(1))\binom{n}{k}, & k > dn(1+\delta_n). \end{cases}$$

Proof. For $k < dn$ note that G_n is a subgraph of M_n and so $N(G_n, k) \leq N(M_n, k)$. Moreover $\mathcal{A}_{\text{pair}}$ is a (G_n, k)-star of size $\binom{n}{k} - \binom{n-2}{k}$.

For $k > dn(1+\delta_n)$ a similar proof to that already given for Theorem 2.1 can be used (for more details see Corollary 3.3 in the next section). $\square$

Bohman et al. [2] asked whether other types of extremal G-intersecting families can occur (apart from families which are either (G, k)-stars or consist of almost all of $\binom{[n]}{k}$). We show that they can by giving a simple example of a graph for which (for appropriate values of k) the extremal family must be of a third type.

Let H_n be the graph with vertex set $[n]$ and edges $\{1, 2\}, \{3, 4\}, \{5, 6\}$. The following family is H_n-intersecting

$$\mathcal{A}_2 = \left\{ A \in \binom{[n]}{k} : A \text{ meets at least two of the three edges of } H_n \right\}.$$

Note that

$$|\mathcal{A}_2| = \binom{n}{k} - 3\binom{n-4}{k} + 2\binom{n-6}{k}.$$

Let $\epsilon > 0$ be small, $n \geq n_0(\epsilon)$ be large and $1 - 2^{-1/2} + \epsilon < k/n < 1/2 - \epsilon$. Since $k/n > 1 - 2^{-1/2} + \epsilon$ it is straightforward to check that $\mathcal{A}_2$ is larger than the largest (H_n, k)-star. (The largest (H_n, k)-star consists of all k-sets meeting a fixed edge of H_n and so has size $\binom{n}{k} - \binom{n-2}{k}$.) Moreover $N(H_n, k) \neq (1 - o(1))\binom{n}{k}$, since $k/n < 1/2 - \epsilon$ implies (by the Erdős–Ko–Rado theorem) that any H_n-intersecting family contains at most $\binom{n-7}{k-1}$ of the sets in $\binom{[n]\setminus[6]}{k}$, so $N(H_n, k) \leq \binom{n}{k} - \binom{n-7}{k}$.

Hence if $\mathcal{B}$ is a k-uniform H_n-intersecting family of maximum size then $\mathcal{B}$ does not contain almost all k-sets and $\mathcal{B}$ is not an (H_n, k)-star. We do not know what form $\mathcal{B}$ can take, only that it must be of some new third type, however $\mathcal{A}_2$ is an obvious candidate extremal family.

3 General k-uniform problem: k large

The conclusion of Theorem 2.1 for k large can be extended to give an analogous result in a more general setting. The exact formulation of this generalization (Theorem 3.1) is rather ugly however we give two natural corollaries (Corollaries 3.2 and 3.3). Recall the definition (1) of a (G, k)-star with centre C from the introduction.

Theorem 3.1 has the following intuitive interpretation: if a largest (G, k)-star contains slightly more than half of all k-sets and there are lots of "independent" (G, k)-stars of this size then the "majority family", consisting of all k-sets belonging to more than half of these (G, k)-stars, contains almost all k-sets. (Idea of proof: a random k-set belongs to any particular largest (G, k)-star with probability $1/2 + \epsilon$ so with high probability it belongs to a majority of them.)

Theorem 3.1. *Let $\{G_n\}_{n=1}^{\infty}$ be a sequence of graphs of order n and $1 \leq k \leq n$. If the following three conditions hold for all n sufficiently large then $N(G_n, k) = (1 - o(1))\binom{n}{k}$.*

(i) There exist isomorphic (G_n, k)-stars: $\mathcal{A}_1^*$, $\mathcal{A}_2^*$, $\ldots, \mathcal{A}_{w_n}^*$ with pairwise disjoint centres $C_1, C_2, \ldots, C_{w_n}$ such that $w_n \to \infty$ as $n \to \infty$.

(ii) The common size of the centres of the $\mathcal{A}_i^*$ is γ_n.

(iii) Each of the $\mathcal{A}_i^*$ has size S_n satisfying

$$S_n \geq \left(\frac{1}{2} + (\gamma_n + 1)\sqrt{\frac{\log w_n}{w_n}}\right)\binom{n}{k}.$$

Let C_n^p denote the pth power of the n-cycle. (That is the graph with vertex set $[n]$ and $i \sim_{C_n^p} j$ iff $1 \leq \mathrm{dist}(i, j) \leq p$, where distance is measured around the cycle.)

Corollary 3.2. *Let $p \geq 1$ be a constant and let α_p be the smallest positive root of*

$$(1 - x)^{2p+1}(1 + px) = 1/2.$$

There exists $\epsilon_{p,n} = o(n)$ such that if $k \geq \alpha_p n + \epsilon_{p,n}$ then $N(C_n^p, k) = (1 - o(1))\binom{n}{k}$. In particular $N(C_n, k) = (1 - o(1))\binom{n}{k}$ for $k > 0.266n$.

Corollary 3.3. *If $r \geq 1$ is a constant and the number of pairwise disjoint r-cliques in G_n is unbounded as $n \to \infty$ then there exists $\epsilon_{r,n} = o(n)$ such that $N(G_n, k) = (1 - o(1))\binom{n}{k}$, for $k > (1 - 2^{-1/r})n + \epsilon_{r,n}$.*

We note that both Corollaries 3.2 and 3.3 could be extended to the case of p, r non-constant but for simplicity we omit these extensions.

Proof of Theorem 3.1. Suppose that $\{G_n\}_{n=1}^{\infty}$, $\mathcal{A}_1^*$, $\mathcal{A}_2^*, \ldots, \mathcal{A}_{w_n}^*$, and $C_1, C_2, \ldots, C_{w_n}$ satisfy conditions (i)–(iii). For $A \in \binom{[n]}{k}$ let

$$\lambda(A) = \#\{i \in [w_n] : A \in \mathcal{A}_i^*\}$$

and define

$$\mathcal{A}_{\mathrm{maj}} = \{A \in \binom{[n]}{k} : \lambda(A) > w_n/2\}.$$

Clearly $\mathcal{A}_{\mathrm{maj}}$ is G_n-intersecting since if $A, B \in \mathcal{A}_{\mathrm{maj}}$ then there exists $i \in [w_n]$ such that $A, B \in \mathcal{A}_i^*$ and $\mathcal{A}_i^*$ is G_n-intersecting (since it is a (G_n, k)-star). We will adapt the proof method of Theorem 2.1 to show that $|\mathcal{A}_{\mathrm{maj}}| = (1 - o(1))\binom{n}{k}$.

Let $A \in \binom{[n]}{k}$ be chosen uniformly at random. For $j \in [n]$ let

$$X_j(A) = \begin{cases} 1, & j \in A, \\ 0, & \text{otherwise.} \end{cases}$$

Since the centres $C_1, C_2, \ldots, C_{w_n}$ are pairwise disjoint and have common size γ_n we may suppose that for $i \in [w_n]$ we have $C_i = [(i-1)\gamma_n + 1, i\gamma_n]$. For $i \in [w_n]$ let $\mathcal{F}_i$ be the σ-algebra generated by $X_1(A), X_2(A), \ldots, X_{i\gamma_n}(A)$. (That is we condition on how A meets the centres of $\mathcal{A}_1^*, \mathcal{A}_2^*, \ldots, \mathcal{A}_i^*$.) Let $Y_0 = \mathbb{E}(w_n - \lambda(A))$ and for $i \in [w_n]$ define $Y_i = \mathbb{E}(w_n - \lambda(A)|\mathcal{F}_i)$. Now $Y_0, Y_1, \ldots, Y_{w_n}$ is a martingale and, since each centre has size γ_n, we have

$$|Y_i - Y_{i-1}| \leq \gamma_n + 1,$$

for $i \in [w_n]$. Moreover

$$Y_0 = w_n \left(1 - \frac{S_n}{\binom{n}{k}}\right) \leq \frac{w_n}{2} - (\gamma_n + 1)\sqrt{w_n \log w_n}.$$

Applying Azuma's inequality (Lemma 2.9) we obtain

$$\begin{aligned} \Pr(A \notin \mathcal{A}_{\text{maj}}) &= \Pr(Y_{w_n} \geq w_n/2) \\ &\leq \Pr(Y_{w_n} \geq Y_0 + (\gamma_n+1)\sqrt{w_n \log w_n}) \\ &\leq \frac{1}{\sqrt{w_n}} \\ &= o(1). \end{aligned}$$

The result now follows. □

Proof of Corollary 3.2. A largest (C_n^p, k)-star, $\mathcal{C}^*$, is given by taking a largest clique K (of order $p+1$) and all $\binom{p+1}{2}$ pairs of vertices $\{i,j\} \in \binom{[n]\setminus K}{2}$ satisfying $i \not\sim_{C_n^p} j$ and $K \subseteq \Gamma^+(\{i,j\})$. Hence

$$|\mathcal{C}^*| = \binom{n}{k} - \binom{n-2p-1}{k} - p\binom{n-2p-2}{k-1}.$$

Note that the centre of $\mathcal{C}^*$ has size $3p+1$. Moreover the number of (C_n^p, k)-stars of maximum size with pairwise disjoint centres is at least $\lfloor n/(3p+1) \rfloor \to \infty$ as $n \to \infty$. (So conditions (i) and (ii) of Theorem 3.1 hold.)

Writing $c = k/n$ we have

$$|\mathcal{C}^*| \geq \left(1 - (1-c)^{2p+1}(1+cp) + O\left(\frac{1}{n}\right)\right)\binom{n}{k}.$$

Hence, for a suitable choice of $\epsilon_{p,n}$ (which can clearly be taken to satisfy $\epsilon_{p,n} = o(n)$), if $k \geq \alpha_p n + \epsilon_{p,n}$ then condition (iii) of Theorem 3.1 also holds and the result follows. □

Proof of Corollary 3.3. This is almost identical to the proof of Corollary 3.2 so we give only a sketch. Let $K_1, K_2, \ldots, K_{w_n}$ be pairwise disjoint r-cliques in G_n, with $w_n \to \infty$ as $n \to \infty$. If

$$\mathcal{A}_i^* = \left\{A \in \binom{[n]}{k} : A \cap K_i \neq \emptyset\right\}$$

then $|\mathcal{A}_i^*| = \binom{n}{k} - \binom{n-r}{k}$. Conditions (i) and (ii) of Theorem 3.1 hold with $\gamma_n = r$. Moreover there exists $\epsilon_{r,n} = o(n)$ such that if $k > (1 - 2^{-1/r})n + \epsilon_{r,n}$ then condition (iii) also holds. The result now follows. □

4 Non-uniform G-intersecting families

The question of how large a non-uniform intersecting family $\mathcal{A} \subseteq 2^{[n]}$ can be is rather easy: $\mathcal{A}$ cannot contain both a set and its complement and so $|\mathcal{A}| \leq 2^{n-1}$, moreover this bound can attained in numerous different ways.

The non-uniform G-intersection problem is also easier to solve than the k-uniform version. For a graph G of order n let

$$N(G) = \max\{|\mathcal{A}| : \mathcal{A} \subseteq 2^{[n]} \text{ is } G\text{-intersecting}\}.$$

The size of the extremal family depends on the matching number, $m(G)$, the size of a largest matching in G.

Theorem 4.1. *If $\{G_n\}_{n=1}^{\infty}$ is a sequence of non-empty graphs of order n with $m(G_n)$ non-decreasing then either $m(G_n) \to \infty$ as $n \to \infty$, in which case $N(G_n) = (1 - o(1))2^n$, or there exists $m \geq 1$ such that $m(G_n) = m$ for all $n \geq n_0$ and*

$$1 - e^{-m/8} \leq \mathbb{P}(Bin(m, 3/4) > m/2) \leq \frac{N(G_n)}{2^n} \leq 1 - 2^{-(2m+1)}. \tag{17}$$

In the latter case both bounds are attainable.

Lemma 4.2. *Let G be a graph of order n with a matching of size $m \geq 1$ then*

$$\frac{N(G)}{2^n} \geq \mathbb{P}(Bin(m, 3/4) > m/2) \geq 1 - e^{-m/8}.$$

Proof. Take a matching M of size m and define

$$\mathcal{M}_{\text{maj}} = \{A \in 2^{[n]} : A \text{ meets} > m/2 \text{ of the edges in } M\}. \tag{18}$$

Note that $\mathcal{M}_{\text{maj}}$ is G-intersecting.

If we select a set $A \in 2^{[n]}$ uniformly at random by choosing each $i \in [n]$ independently with probability $1/2$ then A meets any edge $e \in M$ independently with probability $3/4$. Hence if $X \sim \text{Bin}(m, 3/4)$ then

$$\mathbb{P}(A \in \mathcal{M}_{\text{maj}}) = \mathbb{P}(X > m/2).$$

Hoeffding's inequality [6] implies that this is at least $1 - e^{-m/8}$ and the result follows. □

Proof of Theorem 4.1. Since $m(G_n)$ is increasing either $m(G_n) \to \infty$ as $n \to \infty$ or there exists $m \geq 1$ such that $m(G_n) = m$ for all $n \geq n_0$. In the former case Lemma 4.2 implies that $N(G_n) = (1 - o(1))2^n$, so suppose that $m(G_n) = m$ for all $n \geq n_0$. Lemma 4.2 now implies that the lower bound in (17) holds. For the upper bound consider a maximal matching in G_n, this contains at most $2m$ vertices. Let $W \subseteq [n]$ be the other vertices of G_n. Since W is the complement of a maximal matching it is an independent set and so if $A, B \subseteq 2^W$ are G_n-intersecting then $A \cap B \neq \emptyset$. Hence if $\mathcal{A} \subseteq 2^{[n]}$ is G_n-intersecting then it contains at most half of the sets from 2^W. The fact that $|W| \geq n - 2m$ yields the upper bound in (17).

Note that if G_n is the union of a clique of order $2m + 1$ and $n - 2m - 1$ isolated vertices then the upper bound in (17) is sharp. (The family of all sets meeting the clique is G_n-intersecting and of the correct size.)

To see that the lower bound in (17) is also attainable requires slightly more work. We claim that if G_n is the union of a matching M on m edges and $n - 2m$ isolated vertices then the family $\mathcal{M}_{\text{maj}}$ defined in (18) is a largest G_n-intersecting family. (We will assume for simplicity that m is odd, if m is even a similar argument will work.)

Let $\mathcal{A} \subseteq 2^{[n]}$ be a G_n-intersecting family of maximum size. Let $E = \{e_1, e_2, \ldots, e_m\}$ be the m edges of the matching and let $V = \{v_1, v_2, \ldots, v_{n-2m}\}$ be the $n - 2m$ isolated vertices. For $A \subseteq E$ and $B \subseteq V$ let

$$\mathcal{S}(A, B) = \{C \in 2^{[n]} : A = \{e \in E : C \cap e \neq \emptyset\} \text{ and } B = \{v \in V : v \in C\}\}.$$

So $\mathcal{S}(A, B)$ contains those sets which meet precisely those edges in A and contain precisely those isolated vertices in B.

First note that if $A \subseteq E$ and $B \subseteq V$ then $|\mathcal{S}(A, B)| = 3^{|A|}$. Secondly if $A \subseteq E$ and $B \subseteq V$ then at most one of $\mathcal{A} \cap \mathcal{S}(A, B)$ and $\mathcal{A} \cap \mathcal{S}(E \backslash A, V \backslash B)$ can be non-empty (otherwise $\mathcal{A}$ is not G_n-intersecting). Moreover the maximality of $\mathcal{A}$ implies that if $\mathcal{A} \cap \mathcal{S}(A, B) \neq \emptyset$ then $\mathcal{S}(A, B) \subseteq \mathcal{A}$. Finally note that if for each $A \subseteq E$ and $B \subseteq V$ we take the larger of $\mathcal{S}(A, B)$ and $\mathcal{S}(E \backslash A, V \backslash B)$ then the resulting family is at least as large as $\mathcal{A}$. However this family is $\mathcal{M}_{\text{maj}}$. □

5 Open problems and conjectures

An analogue of Theorem 2.1 should surely hold when $G_{r,n} = K_1 \dot{\cup} K_2 \dot{\cup} \cdots \dot{\cup} K_t$ is the disjoint union of r-cliques, where $r > 2$ is a constant and $n = rt$. Indeed by Corollary 3.3 we have $N(G_{r,n}, k) = (1 - o(1))\binom{n}{k}$ for $k > d_r n(1 + o(1))$ (where $d_r = 1 - 2^{-1/r}$). Moreover if $k < d_r n(1 - o(1))$ then we can prove that $N(G_{r,n}, k) = (1 + o(1))\left(\binom{n}{k} - \binom{n-r}{k}\right)$. However an exact version should hold so we make the following conjecture.

Conjecture 5.1. *If $r > 2$ is a constant, $G_{r,n}$ is a disjoint union of r-cliques and $d_r = 1 - 2^{-1/r}$ then there exists $\delta_{r,n} = o(1)$ such that*

$$N(G_{r,n}, k) = \begin{cases} \binom{n}{k} - \binom{n-r}{k}, & k < d_r n(1 - \delta_{r,n}), \\ (1 - o(1))\binom{n}{k}, & k > d_r n(1 + \delta_{r,n}). \end{cases}$$

Moreover the extremal families are unique up to isomorphism.

Since there is a small range of values of k for which Theorem 2.1 fails to determine $N(M_n, k)$ we ask the following obvious question.

Question 5.2. *Is* $N(M_n, k) = \max\{|\mathcal{A}_{pair}|, |\mathcal{A}_{maj}|\}$ *for all values of* k *and* n*?*

Bohman et al. [2] made the following conjecture concerning the cycle.

Conjecture 5.3 (Bohman et al. [2])**.** *There is a constant* c *such that for any fixed* $\epsilon > 0$

$$N(C_n, k) = \begin{cases} \binom{n}{k} - \binom{n-2}{k} + \binom{n-4}{k-2}, & k < (c - \epsilon)n, \\ (1 - o(1))\binom{n}{k}, & k > (c + \epsilon)n. \end{cases}$$

Given our result for cycles (Corollary 3.2) we make the following conjecture.

Conjecture 5.4. *Conjecture 5.3 is true with* $c = 0.266\ldots$*, the smallest positive root of* $(1-x)^3(1+x) = 1/2$.

Given our example showing that there exist graphs and values of k for which the extremal k-uniform G-intersecting families are neither (G, k)-stars nor almost all of $\binom{[n]}{k}$ we pose the following question.

Question 5.5. *Is it true that for any graph* G *and* $1 \le k \le n$*, there exist* (G, k)*-stars* $\mathcal{A}_1^*, \ldots, \mathcal{A}_t^*$ *such that* $N(G, k) = |\mathcal{A}_{maj}|$*? Where for* t *odd*

$$\mathcal{A}_{maj} = \left\{ A \in \binom{[n]}{k} : A \text{ belongs to } > t/2 \text{ of the } \mathcal{A}_i^* \right\}$$

and for t *even we extend this family to include as many* k*-sets as possible that belong to exactly* $t/2$ *of the* $\mathcal{A}_i^*$.

We note that a result of Erdős, Frankl and Katona [4] implies that a positive answer to this question would yield a positive answer to Question 5.2.

References

[1] Azuma, K. (1967) Weighted sums of certain dependent random variables. *Tôhoku Math. J.* **19** 357-367.

[2] Bohman, T., Frieze, A., Ruszinkó, M. and Thoma, L. (2000) G-intersecting families. *Combin. Probab. and Comput.* **10** 376–384.

[3] Bohman, T. and Martin, R. (2003) A note on G-intersecting families. *Discrete. Math.* **260** 183–188.

[4] Erdős, P. L., Frankl, P. and Katona, G. O. H. (1985) Extremal hypergraph problems and convex hulls. *Combinatorica* **5** 11–26.

[5] Erdős, P., Ko, C. and Rado, R. (1961) Intersection theorems for systems of finite sets. *Q. J. Math. Oxford Ser. 2* **12** 313–320.

[6] Hoeffding, W., (1963) Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** (301) 13–30.