

Some Intersection Theorems for Ordered Sets and Graphs

F. R. K. CHUNG* AND R. L. GRAHAM

*AT&T Bell Laboratories, Murray Hill, New Jersey 07974 and
Bell Communications Research, Morristown, New Jersey

P. FRANKL

C.N.R.S., Paris, France

AND

J. B. SHEARER¹

University of California, Berkeley, California

Communicated by the Managing Editors

Received May 22, 1984

A classical topic in combinatorics is the study of problems of the following type: What are the *maximum* families \mathbf{F} of subsets of a finite set with the property that the intersection of any two sets in the family satisfies some specified condition?

Typical restrictions on the intersections $F \cap F'$ of any F and F' in \mathbf{F} are:

- (i) $F \cap F' \neq \emptyset$, where all $F \in \mathbf{F}$ have k elements (Erdős, Ko, and Rado (1961)).
- (ii) $|F \cap F'| \geq j$ (Katona (1964)).

In this paper, we consider the following general question: For a given family \mathbf{B} of subsets of $[n] = \{1, 2, \dots, n\}$, what is the largest family \mathbf{F} of subsets of $[n]$ satisfying

$$F, F' \in \mathbf{F} \Rightarrow F \cap F' \supseteq B \quad \text{for some } B \in \mathbf{B}.$$

Of particular interest are those \mathbf{B} for which the maximum families consist of so-called "kernel systems," i.e., the family of all *supersets* of some fixed set in \mathbf{B} . For example, we show that the set of all (cyclic) translates of a block of consecutive integers in $[n]$ is such a family. It turns out rather unexpectedly that many of the results we obtain here depend strongly on properties of the well-known *entropy* function (from information theory).

¹ Current address: IBM Research, PO Box 218, Yorktown Heights, New York 10598.

I. INTRODUCTION

A classical topic in combinatorics is the study of questions of the following type: What are the *maximum* families \mathbf{F} of subsets of a finite set with the property that the intersection of any two sets in the family satisfies some specified condition?

Typical restrictions on the intersections based on F and F' in \mathbf{F} are:

- (i) $\bar{F} \cap F' \neq \emptyset$, where \bar{F} denotes the complement of F [16];
- (ii) $F \cap F' \neq \emptyset$, where all $F \in \mathbf{F}$ have k elements [3];
- (iii) $|F \cap F'| \geq j$ [8].

Good surveys of our current state of knowledge in this area can be found in [6, 7, 9, 17], in addition to the results in [5, 12, 13, 14, 18].

In this note we investigate the following question: For a given family \mathbf{B} of subsets of $[n] := \{1, 2, \dots, n\}$, what is the largest family \mathbf{F} of subsets of $[n]$ satisfying:

$$F, F' \in \mathbf{F} \Rightarrow F \cap F' \supseteq B \quad \text{for some } B \in \mathbf{B}. \quad (1)$$

In particular, let $v(\mathbf{B})$ denote the cardinality of the largest family \mathbf{F} satisfying (1).

An Easy Example

As a prelude to the general results, we first consider a simple special case. For $\mathbf{B} = \mathbf{B}_2$ we take the set of all pairs $\{i, i+1\}$, $1 \leq i < n$. For the family \mathbf{B}_2 we prove

$$v(\mathbf{B}_2) = 2^{n-2}. \quad (2)$$

Proof of (2): Define S_i , $i = 1, 2$, by

$$S_i := \{j \in [n] : j \equiv i \pmod{2}\}. \quad (3)$$

Observe that for all i and all $B \in \mathbf{B}$

$$S_i \cap B \neq \emptyset. \quad (4)$$

Suppose $\mathbf{F} \subseteq 2^{[n]}$ satisfies (1). Define the induced families $\mathbf{F}(S_i)$ by

$$\mathbf{F}(S_i) := \{F \cap S_i : F \in \mathbf{F}\}, \quad i = 1, 2. \quad (5)$$

Note that if $G, G' \in \mathbf{F}(S_i)$ then

$$\begin{aligned} G \cap G' &= (F \cap S_i) \cap (F' \cap S_i) \quad \text{for some } F, F' \in \mathbf{F} \\ &= F \cap F' \cap S_i \neq \emptyset \end{aligned} \quad (6)$$

since $F \cap F' \supseteq B'$ for some $B' \in \mathbf{B}$ and by construction $S_i \cap B \neq \emptyset$ for every $B \in \mathbf{B}$. Thus, for $i = 1, 2$, $\mathbf{F}(S_i)$ is a family of subsets of S_i with the property that no two sets in $\mathbf{F}(S_i)$ are disjoint. This implies that

$$|\mathbf{F}(S_i)| \leq \frac{1}{2} \cdot 2^{|S_i|} \tag{7}$$

since we cannot have a set X and its complement $S_i - X$ both in $\mathbf{F}(S_i)$. Since any set $F \in \mathbf{F}$ is determined by its intersections $F \cap S_i$, $i = 1, 2$, then by (7)

$$|\mathbf{F}| \leq \frac{1}{2} \cdot 2^{|S_1|} \cdot \frac{1}{2} \cdot 2^{|S_2|} = \frac{1}{4} \cdot 2^{|S_1| + |S_2|} = 2^{n-2}. \tag{8}$$

On the other hand, for the family \mathbf{F}' given by $\mathbf{F}' = \{X \subseteq [n]: \{1, 2\} \subseteq X\}$, we have

$$F \cap F' \subseteq \{1, 2\} \in \mathbf{B} \quad \text{for all } F, F' \in \mathbf{F}'$$

and

$$|\mathbf{F}'| = 2^{n-2}.$$

This proves (2). ■

Note that the content of (2) is just that no family satisfying (1) for B_2 can have more sets than can be achieved in a trivial way, i.e., by taking all subsets of $[n]$ containing a fixed $B_0 \in \mathbf{B}$. In general, we call such a family a *kernel system* with kernel B_0 . Of course, (2) does not imply that every maximum family \mathbf{F} is a kernel system.

In what follows, we will be especially interested in those families \mathbf{B} for which $v(\mathbf{B})$ is attained by kernel systems. This seems to be true, for example, for any family \mathbf{B} formed by taking the (cyclic) translates of a fixed set in $[n]$ (although we do not prove this).

II. PARTITIONS OF $[n]$

Although we study set *intersections* here, it is sometimes useful to consider the following variation of set intersection, namely, the complement of the symmetric difference of two sets, defined for $X, Y \subseteq [n]$ by

$$X \nabla Y := (X \cap Y) \cup (\bar{X} \cap \bar{Y}) = \overline{X \Delta Y}$$

where $\bar{X} = [n] - X$. For a given family \mathbf{B} of subsets of $[n]$, let $\bar{v}(\mathbf{B})$ denote the cardinality of the largest family \mathbf{F} satisfying

$$F, F' \in \mathbf{F} \Rightarrow F \nabla F' \supseteq B \quad \text{for some } B \in \mathbf{B}.$$

Obviously $v(\mathbf{B}) \leq \bar{v}(\mathbf{B})$.

Slightly less obvious is the following.

FACT 1.

$$v(\mathbf{B}) = \bar{v}(\mathbf{B}) \text{ for all } \mathbf{B}. \tag{9}$$

Sketch of proof. Assume \mathbf{F} is a maximum ∇ -family for \mathbf{B} , i.e., $|\mathbf{F}| = \bar{v}(\mathbf{B})$. Select, if possible, some element $t \in [n]$ so that for some $F \in \mathbf{F}$, $F \cup \{t\} \notin \mathbf{F}$. Replace all such $F \in \mathbf{F}$ (simultaneously) by $F \cup \{t\}$, forming a new family \mathbf{F}' . It is easy to check that \mathbf{F}' is also a ∇ -family for \mathbf{B} , and $|\mathbf{F}'| = |\mathbf{F}|$. Continue this process as long as possible, finally forming the family \mathbf{F}^* , which has the property that for any $F \in \mathbf{F}'$, if $t \notin F$ then $F \cup \{t\} \in \mathbf{F}^*$. Thus, \mathbf{F}^* is an upper ideal in the lattice of subsets $2^{[n]}$, i.e., $[n] \supseteq G \supset F \in \mathbf{F}^*$ implies $G \in \mathbf{F}^*$. It now follows easily that \mathbf{F}^* is in fact an \cap -family for \mathbf{B} , i.e., $F, F' \in \mathbf{F}^*$ implies $F \cap F' \supseteq B$ for some $B \in \mathbf{B}$. Since $|\mathbf{F}^*| = |\mathbf{F}| = \bar{v}(\mathbf{B})$ then we have $v(\mathbf{B}) \geq |\mathbf{B}|$ which implies (9). ■

THEOREM 1. *Suppose $[n] = S_1 \cup \dots \cup S_k$ is a partition of $[n]$ into k non-empty subsets. For $X \subseteq [n]$, define $f(X) = \{i: S_i \cap X \neq \emptyset\}$. Let \mathbf{B} be a family of subsets of $[n]$ and define $\mathbf{B}^* = \{f(X): X \in \mathbf{B}\} \subseteq 2^{[k]}$. Then we have*

$$v(\mathbf{B}) \leq v(\mathbf{B}^*) 2^{n-k}. \tag{10}$$

Proof. By Fact 1, it is enough to prove

$$\bar{v}(\mathbf{B}) \leq \bar{v}(\mathbf{B}^*) 2^{n-k}. \tag{10'}$$

Let \mathbf{F} be a ∇ -family for \mathbf{B} , i.e., $F, F' \in \mathbf{F}$ implies $F \nabla F' \supseteq B$ for some $B \in \mathbf{B}$. Also, let \mathbf{W} denote the subspace of $2^{[n]}$ (considered as an n -dimensional vector space under the operation Δ) generated by the S_i . Partition $2^{[n]}$ into cosets $C_i \Delta \mathbf{W}$, $1 \leq i \leq 2^{n-k}$. It will suffice to show that each coset $C \Delta \mathbf{W}$ contains at most $\bar{v}(\mathbf{B}^*)$ elements of \mathbf{F} . Since $(X \Delta C) \nabla (Y \Delta C) = X \nabla Y$, it suffices to prove that \mathbf{W} contains at most $\bar{v}(\mathbf{B}^*)$ elements of \mathbf{F} . Note that f is a one-to-one map of \mathbf{W} to $2^{[k]}$ and it is easily checked that $f(X \nabla Y) = f(X) \nabla f(Y)$. Hence, for $F, F' \in \mathbf{F} \cap \mathbf{W}$, we have

$$f(F) \nabla f(F') = f(F \nabla F') \supseteq f(B) \in \mathbf{B}^*$$

for some $B \in \mathbf{B}$. Therefore, \mathbf{W} contains at most $\bar{v}(\mathbf{B}^*)$ elements of \mathbf{F} and Theorem 1 is proved. ■

As an immediate consequence of Theorem 1, we have the following result, which has also been obtained independently by Faudree, Schelp, and Sós [4].

THEOREM 2. *Suppose $[n] = S_1 \cup \dots \cup S_k$ is a partition of $[n]$ into k non-empty sets, and $\mathbf{B} \subseteq 2^{[n]}$ is a family with the property that for some j , $1 \leq j \leq k$, each $B \in \mathbf{B}$ intersects at least j of the S_i , $1 \leq i \leq k$. Then*

$$v(\mathbf{B}) \leq 2^{n-k} g(k, j) \tag{11}$$

where

$$g(k, j) = \begin{cases} \sum_{t \geq v} \binom{k}{t} & \text{if } k + j = 2v, \\ \sum_{t \geq v} \binom{k}{t} + \binom{k-1}{v-1} & \text{if } k + j = 2v - 1. \end{cases}$$

Proof. By Fact 1 and Theorem 1 we have

$$v(\mathbf{B}) \leq v(\mathbf{B}^*) 2^{n-k}.$$

Since \mathbf{B}^* is a family of subsets of $[k]$ each containing at least j elements, then a result of Kleitman [10] (also see Ahlswede and Katona [1]) implies $v(\mathbf{B}^*) \leq g(k, j)$. This proves Theorem 2. ■

In order to apply Theorem 2 to a particular family \mathbf{B} , we need to choose a suitable partition $[n] = \bigcup_{i=1}^k S_i$ (which determines some maximal value of j associated with it). It is always possible to use trivial partitions and indeed, these are sometimes optimal. For example, for $[n] = S_1$ we have $k = 1$, $j = 1$, $g(k, j) = 1$, and so,

$$v(\mathbf{B}) \leq 2^{n-1}$$

for any family \mathbf{B} (which does not contain \emptyset). Of course, for $\mathbf{B} = \{\{1\}\}$, for example, the family $\mathbf{F} = \{X \subseteq [n] : 1 \in X\}$ shows that this bound can be achieved.

On the other hand, suppose we take for \mathbf{B} the family of all j -element subsets of $[n]$. For the (maximum) partition $[n] = \bigcup_{i=1}^n S_i$ with $S_i = \{i\}$, the condition that $F \cap F' \supseteq B$ for some $B \in \mathbf{B}$ is equivalent to $|F \cap F'| \geq j$, i.e., $F \cap F'$ intersects at least j of the S_i . In this case, it follows that

$$v(\mathbf{B}) \leq g(n, j). \tag{12}$$

In fact, a theorem of Katona in [6] shows that we actually have equality in this case as well.

For any family \mathbf{B} , if $\mu(\mathbf{B})$ denotes the cardinality of a minimum set B_0 in \mathbf{B} then by forming a maximum kernel system with kernel B_0 , we have

$$v(\mathbf{B}) \geq 2^{n-\mu(\mathbf{B})}. \tag{13}$$

In order to obtain the exact value of $v(\mathbf{B})$ using (11) and (13) it is necessary that

$$g(k, j) = 2^{k - \mu(\mathbf{B})}. \quad (14)$$

As an illustration of (14) let $\mathbf{B}(t)$ denote the family of n t -sets of $[n]$ formed by choosing (cyclically) t consecutive elements of Z_n . We claim that if $n \geq t^2 - t$ then it is always possible to partition $[n]$ into $t + 1$ subsets S_i , $1 \leq i \leq t + 1$, so that the distance (in the corresponding n -cycle C_n) between any s and $s' \in S_i$ is at least t . (An easy way to do this is to write $n = ut + v$, $0 \leq v < t$, write down the string $1, 2, \dots, t, 1, 2, \dots, t, \dots, 1, 2, \dots, t$ of u copies of $1, 2, \dots, t$, and then "insert" v copies of $t + 1$ which are all at distance at least t from one another; this now defines a partition of $[n]$ into $t + 1$ subsets with the desired property.) Since any $B \in \mathbf{B}(t)$ intersects at least t of the $t + 1$ S_i 's then the appropriate values of k and j to use in (11) are $k = t + 1$, $j = t$. However, since $\mu(\mathbf{B}(t)) = t$ then

$$g(t + 1, t) = 2 = 2^{t+1-t}$$

i.e., (14) holds, and consequently

$$v(\mathbf{B}(t)) = 2^{n-t}$$

when $n \geq t^2 - t$. In the next section we will extend this to all values of $n > t$.

III. ON TRANSLATES OF A BLOCK

In this section we will show that for any $t < n$, the collection $\mathbf{B}(t) \subseteq 2^{[n]}$ consisting of a kernel system is the largest intersection family for $\mathbf{B}(t)$ which consists of all cyclic translates of t consecutive numbers. First we will make some easy observations.

FACT 2. *Let $r \leq n/2$. Let X be a subset of the n -cycle C_n such that for $u, v \in X$, the distance between u and v in C_n is no more than $r - 1$. Then $|X| \leq r$.*

Proof. Note that each vertex v in X excludes an interval, denoted by $I(v)$, of length $n + 1 - 2r \geq 1$. We will encounter the $I(v)$, $v \in X$, in the following order. Choose a fixed vertex $v = v_1$. In general, v_i is defined to be the vertex in $X - \{v_i, \dots, v_{i-1}\}$ closest to $\{v_1, \dots, v_{i-1}\}$ (in case of a tie, choose arbitrarily). Now $I(v_1)$ eliminates $n + 1 - 2r$ vertices from C_n . Each additional $I(v_i)$ eliminates at least one more vertex from C_n . Hence the total number of excluded vertices is at least $n + 1 - 2r + |X| - 1$. These

together with the $|X|$ points in X , total at most n . Therefore, $n + 2|X| - 2r \leq n$, i.e., $|X| \leq r$. ■

THEOREM 3. *Suppose $t < n \leq 2t$. Let $\mathbf{B}'(t)$ consist of the cyclic translates of both $\{1, 2, \dots, t\} \pmod n$ together with $\{1, 2, \dots, t\} \pmod{(n-1)}$. Let \mathbf{F} be a family of subsets of $[n]$ with the property that $F, F' \in \mathbf{F} \Rightarrow F \nabla F' \supseteq B$ for some $B \in \mathbf{B}'(t)$. Then we have $|\mathbf{F}| \leq 2^{n-t}$.*

Proof. Since $(X \Delta Z) \nabla (Y \Delta Z) = X \nabla Y$, we may consider $\mathbf{F}' = \{F \Delta F_0 : F' \in \mathbf{F}'\}$ for a fixed subset F_0 in \mathbf{F} . Thus, \mathbf{F}' contains the empty set and $F, F' \in \mathbf{F}' \Rightarrow F \nabla F' \supseteq B$ for some $B \in \mathbf{B}'(t)$. Furthermore $|\mathbf{F}| = |\mathbf{F}'|$ since $F' \neq F''$ if and only if $F_0 \Delta F' \neq F_0 \Delta F''$. It suffices to show $|\mathbf{F}'| \leq 2^{n-t}$. Let U denote the set $\bigcup_{F \in \mathbf{F}'} F = \{x : x \in F \in \mathbf{F}'\}$. Suppose $i, j \in U$, $i, j \neq n$, and $[n] \pmod n$ is viewed as an n -cycle. Then we claim the distance between i and j is at most $n - t - 1$. Assume the contrary. First, suppose i and j both are in $F \in \mathbf{F}'$. Then $F \nabla \emptyset = \bar{F}$ does not contain i and j and cannot contain a cyclic translate of $\{1, \dots, t\} \pmod n$ or $\pmod{(n-1)}$, which is a contradiction. Suppose i and j are in different subsets F, F' in \mathbf{F} . Then again we have $i, j \notin F \nabla F'$ and $F \nabla F'$ cannot contain a cyclic translate of $\{1, \dots, t\} \pmod n$ or $\pmod{(n-1)}$. Hence, by Fact 2, U contains at most $n - t$ elements of $[n - 1]$ or $n - t + 1$ elements of $[n]$. Clearly, $\mathbf{F}' \subseteq 2^U$. Hence if $|U| \leq n - t$, then $|\mathbf{F}'| \leq 2^{n-t}$. Suppose $|U| = n - t + 1$. Let X be a subset of U and $X' = U - X$. Since $(X \nabla X') \cap U = \emptyset$, then $|X \nabla X'| \leq n - |U| \leq t - 1$. Therefore $X \nabla X'$ cannot contain a translate of $\{1, \dots, t\}$ and X, X' cannot both be in U . Hence \mathbf{F}' contains at most half of the subsets in 2^U , i.e., $|\mathbf{F}'| \leq \frac{1}{2} \cdot 2^{n-t+1} = 2^{n-t}$, which completes the proof of Theorem 3. ■

THEOREM 4. *Let \mathbf{F} be a family of subsets of $[n]$ such that $F, F' \in \mathbf{F} \Rightarrow F \nabla F'$ contains some cyclic translate of $\{1, \dots, t\}$. Then $|\mathbf{F}| \leq 2^{n-t}$.*

Proof. By Theorem 3 we only have to consider the case that $n > 2t$. We can write any n as $im + j(m-1)$ for some m , $t < m \leq 2t$, where i, j are non-negative and i is nonzero. Partition $[n]$ into m subsets S_i , $1 \leq i \leq m$, so that the distance between any s and $s' \in S_i$ is at least $m - 1$. Using Theorem 1 we have $v(\mathbf{B}(t)) \leq v(\mathbf{B}^*(t))2^{n-m}$. Theorem 3 then implies $v(\mathbf{B}^*(t)) \leq 2^{m-t}$. Therefore we have $v(\mathbf{B}(t)) \leq 2^{n-t}$ as desired. ■

As an immediate consequence we have the following:

THEOREM 5. *Let \mathbf{F} be a family of subsets of $[n]$ such that $F, F' \in \mathbf{F} \Rightarrow F \cap F'$ contains some cyclic translate of $\{1, \dots, t\}$. Then $|\mathbf{F}| \leq 2^{n-t}$.*

We remark that the kernel system formed by all supersets of $\{1, \dots, t\}$ has 2^{n-t} subsets and hence is a largest possible family.

IV. ON TRANSLATES OF A FIXED SET

We have shown that kernel systems form the best intersection families when \mathbf{B} consists of all the (cyclic) translates of $\{1, 2, \dots, t\}$. It appears that this may hold much more generally.

Conjecture 1. If $\mathbf{B}(X)$ consists of the set of all the cyclic translates of a fixed set $X \subseteq [n]$ then

$$v(\mathbf{B}(X)) = 2^{n-|X|}. \quad (15)$$

Of course a kernel system with kernel X shows that $v(\mathbf{B}(X))$ is at least as large as $2^{n-|X|}$. Although we could not prove this conjecture, the following results provide some evidence in support of the conjecture.

Let $\mathbf{B}_n(X)$ denote the set of all n cyclic translates of X in $[n]$ and let $\mathbf{B}_n^*(X)$ denote the subset of all translates of X . It follows immediately that

$$2^{n-|X|} \leq v(\mathbf{B}_n^*(X)) \leq v(\mathbf{B}_n(X)). \quad (16)$$

Since $v(\mathbf{B}_{n+1}^*(X)) \geq 2v(\mathbf{B}_n^*(X))$, $v(\mathbf{B}_n^*(X))/2^n$ is non-decreasing in n . Consequently,

$$r^*(X) := \lim_{n \rightarrow \infty} \frac{v(\mathbf{B}_n^*(X))}{2^n} \quad \text{exists.} \quad (17)$$

If X is a block of t consecutive integers, then $r^*(X) = 2^{-t}$. We will prove the following:

THEOREM 6.

$$r(X) := \lim_{n \rightarrow \infty} \frac{v(\mathbf{B}_n(X))}{2^n} \quad \text{exists}$$

and

$$r(X) = r^*(X).$$

Proof. From (16) and (17) we have $v(\mathbf{B}_n(X))/2^n \geq v(\mathbf{B}_n^*(X))/2^n$ and $\lim_{n \rightarrow \infty} v(\mathbf{B}_n(X))/2^n = r^*(X)$. Hence, it clearly suffices to show that for any $\varepsilon > 0$ there exists n_0 so that for all $n > n_0$ we have

$$\frac{v(\mathbf{B}_n(X))}{2^n} \leq r^*(X) + \varepsilon.$$

To prove this, it is enough to show for an intersection family \mathbf{F} , we can find a set H of h consecutive integers, where $X \subseteq [h]$, such that

$$|\{F \in \mathbf{F} : F \cap H = \emptyset\}| \geq |\mathbf{F}|/2^h(1 + \varepsilon). \tag{18}$$

To see this, note that $|\{F \in \mathbf{F} : F \cap H = \emptyset\}| \leq v(\mathbf{B}_{n-h}^*(X))$. Combining this with (18) we get

$$\frac{|\mathbf{F}|}{2^n} \leq (1 + \varepsilon) \frac{v(\mathbf{B}_{n-h}^*(X))}{2^{n-h}} \leq r^*(X) + \varepsilon'. \tag{19}$$

We only have to consider \mathbf{F} with

$$|\mathbf{F}| \geq 2^{n-h}. \tag{20}$$

Now we partition $[n]$ into $m = \lceil n/h \rceil$ blocks, i.e., $[n] = S_1 \cup S_2 \cup \dots \cup S_m$, where $|S_m| \leq h$ and $S_i, i \neq m$, is a set of h consecutive numbers. We consider a random variable X assuming values in \mathbf{F} so that each element of \mathbf{F} is equally likely. For $1 \leq i \leq m$, let $X_i = X \cap S_i$ be the associated random variable taking values in $F_i = \{F \cap S_i : F \in \mathbf{F}\}$. We consider the entropy (see [11])

$$H(X) = \sum_F -p_F \log_2 p_F = \log_2 |\mathbf{F}|,$$

where $p_F := \text{Prob}(X = F)$ and the sum is taken over all $F \in \mathbf{F}$. Since X_1, \dots, X_m determine X , we have

$$H(X) \leq \sum_{i=1}^m H(X_i)$$

which with (20) implies

$$n - h \leq \sum_{i=1}^m H(X_i)$$

i.e.,

$$\sum_{i=1}^m (|S_i| - H(X_i)) \leq h.$$

Therefore there exists an i , say $i = 1$, such that

$$|S_1| - H(X_1) \leq \frac{h}{m-1} < \frac{h^2}{n-2h}. \tag{21}$$

Suppose $\text{Prob}(X_1 = \emptyset) < 1/(1 + \varepsilon)2^h$. Then there exists $\delta = \delta(\varepsilon) > 0$, such that

$$H(X_1) < |S_1| - \delta.$$

Therefore we have $h^2/(n - 2h) \geq \delta$ which contradicts the fact that $n > 2h + h^2/\delta$ for n sufficiently large. Thus, $\text{Prob}(X_1 = \emptyset) \geq 1/(1 + \varepsilon)2^h$ and (19) holds. This completes the proof of Theorem 6. ■

Let $X + i$ denote the set $\{x + i \pmod n : x \in X\}$. We have the following.

THEOREM 7. *Suppose $X \subseteq [n]$ satisfies $|X \cup (X + i)| > |X| + \log_2 \binom{n}{2}$ for all $1 \leq i < n$. Then $v(\mathbf{B}(X)) = 2^{n-|X|}$ where equality holds only for kernel systems with kernel $X + j$, for some j .*

Proof. Let $\mathbf{F} \subseteq 2^{[n]}$ be a family of sets such that for any $F, F' \in \mathbf{F}$, $F \cap F'$ contains $X + i$ for some i . We distinguish two cases:

(i) There exists $F \in \mathbf{F}$ such that F contains only one translated copy, say $X + i$, of X . Then $X + i \subseteq F \cap F'$ holds for all $F' \in \mathbf{F}$, i.e., \mathbf{F} is contained in the kernel system $\{F \subseteq [n] : X + i \subseteq F\}$, which has size $2^{n-|X|}$.

(ii) For every $F \in \mathbf{F}$ there are at least two different numbers i, j , $1 \leq i < j \leq n$ such that $(X + i) \subset F$, $(X + j) \subset F$ hold. Since there are only $\binom{n}{2}$ choices for (i, j) there is a particular choice, say k, l , such that $(X + k) \cup (X + l) \subseteq F$ holds for at least $|\mathbf{F}|/\binom{n}{2}$ sets $F \in \mathbf{F}$.

However, $|((X + k) \cup (X + l))| = |X \cup (X + (l - k))| > |X| + \log_2 \binom{n}{2}$, which means that

$$|\{F \subseteq [n] : ((X + k) \cup (X + l)) \subseteq F\}| < 2^{n-|X| - \log_2 \binom{n}{2}} = 2^{n-|X|}/\binom{n}{2}.$$

Consequently $|\mathbf{F}| < \binom{n}{2} 2^{n-|X|}/\binom{n}{2} = 2^{n-|X|}$ holds and Theorem 7 is proved. ■

If $c > 2$ is a constant and $c \log_2 n < t < n - c \log_2 n$ then for almost all t -element subsets X of $[n]$, the assumption of Theorem 7 can be verified. Thus we have:

COROLLARY. *Given c and t satisfying $c > 2$, $c \log_2 n < t < n - c \log_2 n$, then for almost all t -subsets X of $[n]$ we have*

$$v(\mathbf{B}(X)) = 2^{n-|X|}.$$

V. A PRODUCT THEOREM

The following result, which seems to be a very useful tool in many extremal problems in combinatorics, was first proved by one of us (JBS) in

1978 (unpublished). A simpler related result was used by Bombieri [2] in connection with a question of J.-P. Serre.

THE PRODUCT THEOREM. *Let S be a finite set and let A_1, \dots, A_m be subsets of S such that every element of S is contained in at least k of A_1, \dots, A_m . Let \mathbf{F} be a collection of subsets of S and let $\mathbf{F}_i = \{F \cap A_i : F \in \mathbf{F}\}$ for $1 \leq i \leq m$. Then we have*

$$|\mathbf{F}|^k \leq \prod_{i=1}^m |\mathbf{F}_i|.$$

Proof. Let X be a random variable assuming values in \mathbf{F} so that each element of \mathbf{F} is equally likely. For $1 \leq i \leq m$, let $X_i = X \cap A_i$ be the associated random variable taking on values in \mathbf{F}_i . We will prove

$$kH(X) \leq \sum_{i=1}^m H(X_i). \quad (22)$$

If $k=1$, then $S = A_1 \cup \dots \cup A_m$. Thus, X_1, \dots, X_m determine X and consequently, $H(X) \leq \sum_{i=1}^m H(X_i)$ as desired. Now assume $k > 1$. Let j denote the minimum number of A_i 's whose union is S . Clearly $1 \leq j \leq m$. We will prove (22) by induction on k and j . If $j=1$, say $A_1 = S$, we have (by induction on k)

$$(k-1)H(X) \leq \sum_{i \neq 1} H(X_i)$$

and consequently

$$kH(X) \leq \sum_{i=1}^m H(X_i).$$

Suppose $j > 1$. We may assume without loss of generality that $A_1 \cup A_2 \cup \dots \cup A_j = S$. Let $A'_1 = A_1 \cup A_2$, $A'_2 = A_1 \cap A_2$. Clearly every element of S is in at least k of $A'_1, A'_2, A_3, \dots, A_m$. By induction on j we have

$$kH(X) \leq \sum_{i \neq 1,2} H(X_i) + H(X') + H(X'')$$

where $X' = X \cap A'_1$ and $X'' = X \cap A'_2$. Since it can be shown (by the convexity of H) that

$$H(X') + H(X'') \leq H(X_1) + H(X_2)$$

then we have $kH(X) \leq \sum_{i=1}^m H(X_i)$.

Now, $H(X) = \log_2 |\mathbf{F}|$ and $H(X_i) \leq \log_2 |\mathbf{F}_i|$. Thus we have

$$|\mathbf{F}|^k \leq \prod_{i=1}^m |\mathbf{F}_i|$$

and the proof is complete. ■

The following inequalities of interest in information theory can be proved in a similar way. We will state these inequalities but omit the proofs.

$$\begin{aligned} H(X, Y, Z) &\leq \frac{1}{2}(H(X, Y) + H(Y, Z) + H(X, Z)) \\ &\leq H(X) + H(Y) + H(Z). \end{aligned}$$

More generally,

$$H(X_1, \dots, X_t) \leq \binom{t-1}{j-1}^{-1} \sum_{\{i_1, \dots, i_j\} \subseteq [t]} H(X_{i_1}, \dots, X_{i_j}).$$

We will now use the Product Theorem to prove two theorems on intersection families of graphs.

THEOREM 8. *Suppose \mathbf{F} is a family of (labelled) subgraphs of the complete graph K_n such that for all $F, F' \in \mathbf{F}$, $F \cap F'$ does not contain any isolated vertices. Then*

$$|\mathbf{F}| \leq 2^{\binom{n}{2} - \frac{n}{2}}.$$

Proof. Choose A_i to be the (spanning) star at vertex v_i and let $E(A_i)$ denote the set of edges of A_i . Clearly every edge is in exactly two of A_1, \dots, A_n . Now $\mathbf{F}_i = \{F \cap A_i : F \in \mathbf{F}\}$ has the intersection property (i), i.e.,

$$(F \cap A_i) \cap (F' \cap A_i) = (F \cap F') \cap A_i \neq \emptyset.$$

Therefore $|\mathbf{F}_i| \leq 2^{|E(A_i)|-1} = 2^{n-2}$ since for any $T \subset A_i$, T and $A_i - T$ cannot both be in \mathbf{F}_i . Using the Product Theorem, we have

$$|\mathbf{F}|^2 \leq \prod_{i=1}^n |\mathbf{F}_i| \leq 2^{n(n-2)}.$$

Therefore

$$|\mathbf{F}| \leq 2^{n(n-2)/2} = 2^{\binom{n}{2} - \frac{n}{2}}$$

which proves Theorem 8. ■

We note that the bound in Theorem 8 is best possible for the case of n even since one such family is a kernel system consisting of all subgraphs of K_n containing a fixed matching.

THEOREM 9. *Suppose \mathbf{F} is a family of (labelled) subgraphs of K_n such that $F \cap F'$ contains a triangle for all $F, F' \in \mathbf{F}$. Then*

$$|\mathbf{F}| \leq 2^{\binom{n}{2} - 2}.$$

Proof. First, suppose n is even. We choose A_i , $1 \leq i \leq \frac{1}{2}\binom{n}{2}$, to be all possible disjoint unions of two complete (labelled) graphs of $n/2$ vertices each. Then $\mathbf{F}_i = \{F \cap A_i \mid F \in \mathbf{F}\}$ has the intersection property (i) since no triangle can be contained in a bipartite graph. Therefore

$$|\mathbf{F}_i| \leq 2^{|E(A_i)| - 1}.$$

Each edge of K_n is in exactly $\binom{n-2}{n/2}$ A_i 's. Therefore by the Product Theorem we have

$$|\mathbf{F}| \binom{n-2}{n/2} \leq 2^{1/2(2\binom{n/2}{2} - 1)} \binom{n}{n/2}$$

i.e.,

$$\begin{aligned} |\mathbf{F}| &\leq 2^{\binom{n}{2} - n(n-1)/n(n/2-1)} \\ &\leq 2^{\binom{n}{2} - 2}. \end{aligned}$$

For the case of n odd, the proof is quite similar and will be omitted. We remark that the largest such family we can find so far is the kernel system of all $2^{\binom{n}{2} - 3}$ graphs which contain a fixed triangle. The above result supplies evidence in favor of the old conjecture of Simonovits and Sós [15].

Conjecture 2. If \mathbf{F} is a family of (labelled) subgraphs of K_n such that for any $F, F' \in \mathbf{F}$, $F \cap F'$ contains a triangle then $|\mathbf{F}| \leq 2^{\binom{n}{2} - 3}$.

Let $G = K(r_1, r_2, r_3)$ denote the complete tripartite graph on the vertex sets R_i of size r_i , $1 \leq i \leq 3$. Suppose \mathbf{F} is a family of (labelled) subgraphs of G such that $F \cap F'$ contains a triangle for all $F, F' \in \mathbf{F}$. One such family is a kernel system of G containing some fixed triangle. Clearly such a family has $2^{r_1 r_2 + r_2 r_3 + r_3 r_1 - 3}$ graphs in it. We will show that no family \mathbf{F} satisfying the hypothesis can have more than this many graphs. To see this, partition the edge set E of G into three classes E_i , $1 \leq i \leq 3$, where E_i denotes the sets of edges which are not incident to a vertex in R_i . It follows from the structure

of G that $F \cap F'$ must intersect every R_i since all triangles do. Thus, by Theorem 1 we have

$$\begin{aligned} |\mathbf{F}| &\leq 2^{|\mathcal{E}| - 3} g(3, 3) \\ &= 2^{r_1 r_2 + r_2 r_3 + r_3 r_1 - 3} \end{aligned}$$

as claimed.

Here is another tantalizing conjecture:

Conjecture 3. Suppose \mathbf{F} is a family of (labelled) subgraphs of K_n such that for any $F, F' \in \mathbf{F}$, $F \cap F'$ contains a path of three edges. Then

$$|\mathbf{F}| \leq 2^{\binom{n}{2} - 3}$$

i.e., kernel systems give the largest possible families.

At present all that is known is that

$$2^{\binom{n}{2} - 3} \leq \max_{\mathbf{F}} |\mathbf{F}| \leq 2^{\binom{n}{2} - 1},$$

the upper bound resulting from the observation that \mathbf{F} cannot contain a graph and its complement. We remark that if we only consider paths of length 2, then it is not difficult to show that $\max_{\mathbf{F}} |\mathbf{F}| = 2^{\binom{n}{2} - 1 + o(1)}$.

Finally, we mention one more (related) conjecture of Simonovits and Sós [15]:

Conjecture 4. If \mathbf{F} is a family of subsets of $[n]$ such that $F, F' \in \mathbf{F} \Rightarrow F \cap F'$ contains a 3-term arithmetic progression, then $|\mathbf{F}| \leq 2^{n-3}$.

Note that this bound, if true, would be best possible, since in this case the kernel system formed by all sets containing a fixed 3-term arithmetic progression has 2^{n-3} sets in it.

REFERENCES

1. R. AHLSEWEDE AND G. O. H. KATONA, Contributions to the geometry of Hamming spaces, *Discrete Math.* **17** (1977), 1–22.
2. E. BOMBIERI, personal communication.
3. P. ERDŐS, C. KO, AND R. RADO, Intersection theorems for systems of finite sets, *Quart. J. Math.* **2** (1961), 313–320.
4. R. FAUDREE, R. SCHELP, AND V. T. SÓS, Some intersection theorems on two-valued functions, to appear.
5. R. L. GRAHAM, M. SIMONOVITS, AND V. T. SÓS, A note on the intersection properties of subsets of integers, *J. Combin. Theory Ser. A* **28** (1980), 106–110.

6. C. GREENE AND D. J. KLEITMAN, Proof techniques in the theory of finite sets, in "Studies in Combinatorics," M.A.A. Studies in Mathematics Vol. 17, (G.-C. Rota, Ed.), pp. 22-79, 1978.
7. A. J. W. HILTON, On ordered set systems and some conjectures related to the Erdős-Ko-Rado theorem and Turán's theorem, *Mathematika* **28** (1981), 54-66.
8. G. KATONA, Intersection theorems for systems of finite sets, *Acta Math. Acad. Sci. Hungar.* **15** (1964), 329-337.
9. G. O. H. KATONA, Extremal problems among subsets of a finites set, in "Combinatorics" (M. Hall and J. H. van Lint, Eds.), Math. Centrum Tracts Vol. 50, pp. 13-42, Math. Centrum, Amsterdam, 1974.
10. D. J. KLEITMAN, On a combinatorial conjecture of Erdős, *J. Combin. Theory* **1** (1966), 209-214.
11. R. J. MCELEECE, "The Theory of Information and Coding," Addison-Wesley, Reading, Mass., 1977.
12. M. SIMONOVITS AND V. T. SÓS, Graph intersection theorems, in "Proc. Colloq. Combinatorics and Graph Theory," Orsay, Paris, 1976, pp. 389-391.
13. M. SIMONOVITS AND V. T. SÓS, Intersection theorems for subsets of integers, *Notices Amer. Math. Soc.* **25** (1978), A-33.
14. M. SIMONOVITS AND V. T. SÓS, Intersections on structures, *Ann. Discrete Math.* **6** (1980), 301-314.
15. M. SIMONOVITS AND V. T. SÓS, personal communication.
16. E. S. SPERNER, Ein Satz über Untermengen einer endlichen Menge, *Math. Z.* **27** (1928), 544-548.
17. M. DEZA AND P. FRANKL, Erdős-Ko-Rado Theorem 22 years later, *SIAM J. Alg. Disc. Math.* **4** (1983), 419-431.
18. P. FRANKL AND Z. FUREDI, Forbidding just one intersection, *J. Combin. Theory Ser. A* **39** (1985), 160-176.