On Attractive and Friendly Sets in Sequence Spaces

R. Ahlswede and L. Khachatrian

Abstract. To a large extent the present work is far from being conclusive, instead, new directions of research in combinatorial extremal theory are started. Also questions concerning generalizations are immediately noticeable.

The incentive came from problems in several fields such as Algebra, Geometry, Probability, Information and Complexity Theory. Like several basic combinatorial problems they may play a role in other fields. For scenarios of interplay we refer also to [9].

1 Introduction: New Problems and Results

A. A New Isoperimetric Problem: Boundaries with Intensity k, a Counterexample to Keane's Conjecture

For $\mathcal{X} = \{0,1\}$ let d be the Hamming distance in $\mathcal{X}^n = \prod_1^n \mathcal{X}$ and let the pair $\mathcal{H}^n = (\mathcal{X}^n, d)$ be the Hamming space.

$$S(x^n) = \{y^n \in \mathcal{X}^n : d(x^n, y^n) = 1\} \qquad (1.1)$$

is the sphere of radius 1 with center $x^n \in \mathcal{X}^n$.

For any set $V \subset \mathcal{X}^n$ we define

$$\Gamma(V) = \{y^n \in \mathcal{X}^n : d(x^n, y^n) \leq 1 \text{ for some } x^n \in V\} \qquad (1.2)$$

and

$$B(V) = \Gamma(V) \smallsetminus V \qquad (1.3)$$

as the (outer) boundary of V. Harper [4] considered

$$b(n, N) = \min_{V \subset \mathcal{X}^n, |V|=N} |B(V)| \qquad (1.4)$$

and established his well–known Isoperimetric Theorem (in graphic language also called "Vertex Isoperimetric Theorem").

Notice that the points in $B(V)$ have distance 1 with at least one point of V. Our generalisation to a boundary of intensity k is

$$B_k(V) = \{y^n \in \mathcal{X}^n \smallsetminus V, |S(y^n) \cap V| \geq k\}. \qquad (1.5)$$

Obviously $B_1(V) = B(V)$.

Problem 1: Determine

$$b_k(n, N) \triangleq \min_{V \subset \mathcal{X}^n : |V| = N} |B_k(V)| \tag{1.6}$$

and the structure of optimal V.

We have not yet solved it, but we introduced related problems 1–4 and solved problem 2 "ratewise" and problem 3, 4 exactly.

However, we made progress on a problem related to problem 1. For an upset $\mathcal{U}$ we consider $B_k(\mathcal{X}^n \setminus \mathcal{U}) = \mathcal{U}_k \triangleq \{x \in \mathcal{U} : |S(x) \cap (\mathcal{X}^n \setminus \mathcal{U})| \geq k\}$.

M. Keane [Oral communication] defined in the eighties the function

$$f(n, k) \triangleq \max\{|\mathcal{U}_k| : \mathcal{U} \text{ upset in } \mathcal{X}^n\} 2^{-n}$$

and conjectured that

$$\lim_{k \to \infty} \left(\sup_n f(k, n) \right) = 0.$$

Observation: For $k \geq \frac{n}{2}$

$$f(k, n) = \binom{n}{k} 2^{-n}.$$

Indeed, by the AZ–identity [5] for any $\mathcal{A} \subset 2^{[n]}$

$$\sum_{X \subset [n]} \frac{W_{\mathcal{A}}(X)}{|X| \binom{n}{|X|}} \equiv 1,$$

where

$$W_{\mathcal{A}}(X) = \left| \bigcap_{X \supset A \in \mathcal{A}} A \right|,$$

and therefore

$$\sum_{|X| \geq k} W_{\mathcal{U}}(X) \leq \max_{\ell \geq k \geq \frac{n}{2}} \ell \binom{n}{\ell} = k \binom{n}{k}$$

or

$$2^n f(k, n) \leq \frac{1}{k} \sum_{|X| \geq k} W_{\mathcal{U}}(X) \leq \binom{n}{k}.$$

This looks reassuring, however, Keane's conjecture is false.

Example 1

$$[n] = \Omega_1 \dot\cup \Omega_2 \cup \cdots \cup \Omega_m, |\Omega_i| = x, m = \frac{n}{x}.$$

$$\mathcal{A} = \left\{ A \in 2^{[n]} : \exists i, \text{ s.t. } \Omega_i \subset A \right\} = \mathcal{U} \left\{ \Omega_1, \ldots, \Omega_{\frac{n}{x}} \right\},$$

$$\mathcal{U}_x = \{ B \in \mathcal{A} : \exists i, \text{ s.t. } \Omega_i \subset B \text{ and } \Omega_j \not\subset B \text{ for all } j \neq i \},$$

and

$$|\mathcal{U}_x| = \frac{n}{x} \cdot (2^x - 1)^{\frac{n}{x} - 1} = g(x, n), \text{ say.}$$

$\max\limits_{n} g(x, n)$ is assumed at $n = x \cdot 2^x - \frac{x}{1+x\log 2}$.

Hence, we may assume $n = n(x) = x \cdot (2^x - 1)$. For this n we consider

$$\frac{g(x, n)}{2^n} = \frac{(2^x - 1)^{2^x - 1}}{2^{x(2^x - 1)}} = \left(\frac{2^x - 1}{2^x}\right)^{2^x - 1} = \frac{1}{\left(1 + \frac{1}{2^x - 1}\right)^{2^x - 1}}$$

and therefore $\lim\limits_{x\to\infty} \frac{g(x,n)}{2^n} = \lim\limits_{x\to\infty} \left(1 + \frac{1}{2^x - 1}\right)^{-(2^x - 1)} = e^{-1}.$ □

B. The Smallest Rich World Problem

Consider

$$I_k(V) = \{y^n \in V : |S(y^n) \cap V| \geq k\} \tag{1.7}$$

and

Problem 2: Determine

$$r_k(n) = \min_{\substack{\phi \neq V = I_k(V) \\ V \subset \mathcal{X}^n}} |V| \left(= \min_{V \subset \mathcal{X}^n,\ \phi \neq V,\ |I_k(V)|=|V|} |V|\right). \tag{1.8}$$

We call an optimal V k–friendly set (or k–best world). We report now the much more general Problem 7 in [6], which was solved "ratewise" in [7].

For $\varphi : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, $\mathcal{X}$ a finite set, define

$$\alpha = \min_{x,y\in\mathcal{X}} \varphi(x, y), \quad \beta = \max_{x,y\in\mathcal{X}} \varphi(x, y), \tag{1.9}$$

and the sum–type function $\varphi_n : \mathcal{X}^n \times \mathcal{X}^n \to \mathbb{R}$, where $\varphi_n(x^n, y^n) = \sum\limits_{t=1}^{n} \varphi(x_t, y_t)$ for $x^n = (x_1, \ldots, x_2)$ and $y^n = (y_1, \ldots, y_2)$.

Now for any closed interval $\mathrm{L} \subset [\alpha, \beta]$, any positive real number ρ, and any positive integer n call a set $S \subset \mathcal{X}^n$ with the property

$$\left|\left\{y^n \in S : \frac{1}{n}\varphi_n(x^n, y^n) \in \mathrm{L}\right\}\right| \geq 2^{n\rho} \text{ for all } x^n \in S \tag{1.10}$$

(n, L, ρ)–good and denote by $N(n, \mathrm{L}, \rho)$ the smallest cardinality of (n, L, ρ)–good sets.

For the set $\{y^n \in S : \frac{1}{n}\varphi_n(x^n, y^n) \in \mathrm{L}\}$ we also write $B(x^n, \mathrm{L}, S)$. In the case $\alpha = 0$ and $\mathrm{L} = [0, \beta]$ it is the intersection of S with a ball with center x^n and φ–radius β.

Inequality (1.10) says that every point in S has $2^{n\rho}$ points in S in its neighbourhood. In this sense S is a "rich world". The definition of $N(n, \mathrm{L}, \rho)$ catches the goal to make the "world small".

One readily can show that $\lim\limits_{n\to\infty} \frac{1}{n} \log N(n, \mathrm{L}, \rho)$ exists, because $N(n_1+n_2, \mathrm{L}, \rho) \geq N(n_1, \mathrm{L}, \rho) \cdot N(n_2, \mathrm{L}, \rho)$. We denote the limit by $\sigma(\mathrm{L}, \rho)$. Its characterisation requires a few concepts.

Let (U, X, Y) be a triple of RV's with values in $\mathcal{U} \times \mathcal{X} \times \mathcal{X}$. We say that (X, Y) is a matching through U, if for conditional entropies

$$H(X|U) = H(Y|U) \text{ and } H(Y|XU) = H(X|YU). \tag{1.11}$$

Finally, we set
$$\mathcal{Q}(\mathrm{L}, \rho) = \{(X, U) : \text{for some } Y (X, Y) \text{ is matched through } U, \mathbb{E}_\varphi(X, Y) \in \mathrm{L},$$
and $H(Y|XU) \geq \rho\}.$ \tag{1.12}

Theorem AC. [7]
$$\sigma(L, \rho) = \min_{(X, U) \in Q(\acute{L}, \rho)} H(X|U).$$

Actually, we can bound the cardinality of $\mathcal{U}$ by $|\mathcal{X}|^2 + 4$.
Furthermore, we can limit the distributions P_{XY} to those with equal marginals.

Remark: Problem 2 relates to the case $\mathcal{X} = \{0, 1\}$, $\mathrm{L} = [k, n]$ and φ_n as Hamming distance.

C. k–Attractive Sets
In another direction we consider

$$\tilde{V}_k = \{y^n \in \mathcal{X}^n : |S(y^n) \cap V| \geq k\}. \tag{1.13}$$

Clearly,
$$\tilde{V}_k = B_k(V) \dot{\cup} I_k(V). \tag{1.14}$$

Analogous to Problem 1 is

Problem 3: Determine

$$a_k(n) \triangleq \min_{\substack{\phi \neq V : |V| \leq |\tilde{V}_k| \\ V \subset \mathcal{X}^n}} |V| \tag{1.15}$$

and the structure of solutions.

An optimal V is called k–attractive set. We call a V with $|V| \leq |\tilde{V}_k|$ a k–admissible set.

Example 2: Define for $k > 1$

$$\mathcal{X}_+^k = \left\{ x^k = (x_1, \ldots, x_k) \in \mathcal{X}^k : \sum_{t=1}^{k} x_t \equiv 0 \mod 2 \right\}$$

and
$$\mathcal{X}_-^k = \mathcal{X}^k \smallsetminus \mathcal{X}_+^k.$$

Notice that for $V = \mathcal{X}_+^k * (0, 0, \ldots, 0) \subset \mathcal{X}^n$ we have $\tilde{V}_k = \mathcal{X}_-^k * (0, 0, \ldots, 0)$ and that V is k–admissible.

Moreover, let us write

$$Z_+^k = \mathcal{X}_+^k * (0, 0, \ldots, 0), Z_-^k = \mathcal{X}_-^k * (0, 0, \ldots, 0) \subset \mathcal{X}^n \tag{1.16}$$

and observe that for any $\pi \in \Sigma_n$, the symmetric group acting on $\{1, 2, \ldots, n\}$, also $\pi \, Z_+^k$ is k–admissible.

Furthermore, for any $y^n \in GF(2)^n$ also $Z_+^k + y^n$ is k–admissible. Quite surprisingly, these sets lead to all k–attractive sets.

Theorem 1. *For $k \geq 2$ the k–attractive subsets of $\mathcal{X}^n$ are of the form $(\pi \, Z_+^k) + y^n$. In particular for a k–attractive subset V of $\mathcal{X}^n$*

$$|V| = |\tilde{V}_k| = 2^{k-1}.$$

For 1–attractive sets

$$|V| = 1, |\tilde{V}_1| = n.$$

D. k–Pairs

We prove Theorem 1 by deriving it from the solution of the somewhat more general Problem 4:

For $A, B \subset \mathcal{X}^n$ we call (A, B) a k–*pair*, if

$$\sum_{x^n \in B} |A \cap S(x^n)| \geq k|B|. \tag{1.17}$$

It is *admissible*, if $|A| \leq |B|$, and it is *optimal*, if

$$|A| = \min_{(A', B')\text{admissible } k\text{–pair}} |A'|.$$

Determine all optimal k–pairs.

Theorem 2. *For $k \geq 2$ the optimal k–pairs are of the form*

$$(\pi \, Z_+^k + y^n, \pi \, Z_-^k + y^n).$$

The implication of Theorem 1 is readily established.

Clearly, if V is k–admissible, then $(V, \tilde{V}_k)$ is an admissible k–pair. The class of $(V, \tilde{V}_k)$ corresponding to attractive V's constitute a subclass of the class of optimal k–pairs. By Theorem 2 and Example 1 these classes are actually equal.

E. Results for Lopsided Sets in Combinatorial Language

Lopsided sets where introduced in [10] in the study of convex sets. There are several equivalent definitions [11]. We use here the terminology of set theory.

Let $\mathcal{S} \subset 2^{[n]}$, $\mathcal{S}^* = 2^{[n]} \smallsetminus \mathcal{S}$. The set $\mathcal{S}$ is called lopsided, if for all $A \in 2^{[n]}$
either there exists a $B \subset [n] \smallsetminus A$ such that for all $C \subset A \quad B \cup C \in \mathcal{C}$
or there exists a $B' \subset A$ such that for all $C' \subset [n] \smallsetminus A \quad B' \cup C' \in \mathcal{S}^*$.

When Levon told me that Andreas Dress asked him to prove Theorem 3 below, which is considered basic for lopsided sets, I told him that he could and should do it in one afternoon, because this might be helpful for his career.

He followed the suggestion, but his simple proof earned him no benefits.

Theorem 3. *For a set $\mathcal{C} \subset 2^{[n]}$ define*

$$\text{Ł}(\mathcal{C}) \triangleq \{Y \subset [n] : \{Y \cap C : C \in \mathcal{C}\} = 2^Y\}.$$

Then

$$|\mathcal{C}| \le |\text{Ł}(\mathcal{C})|. \tag{1.18}$$

Obviously $\text{Ł}(\mathcal{C})$ is a downset for every $\mathcal{C} \subset 2^{[n]}$ and if $\mathcal{C}$ is a downset, then $\text{Ł}(\mathcal{C}) = \mathcal{C}$. So for downsets there is equality in (1.18).

Recall now the standard push-down operation:

given $\mathcal{C} \subset 2^{[n]}$ and $x \in [n]$ for $A \in \mathcal{C}$

$$T(A, x) = \begin{cases} A \smallsetminus \{x\}, & \text{if } x \in A \text{ and } A \smallsetminus \{x\} \notin \mathcal{C} \\ A & \text{otherwise} \end{cases} \tag{1.19}$$

$$T(\mathcal{C}, x) = \{T(A, x) : A \in \mathcal{C}\}. \tag{1.20}$$

Now readily verify

$$|T(\mathcal{C}, x)| = |\mathcal{C}| \text{ and } \text{Ł}(T(\mathcal{C}, x)) \subset \text{Ł}(\mathcal{C}). \tag{1.21}$$

After finitely many, say m, push–down operations $T(\mathcal{C}, x_1), T((\mathcal{C}, x_1), x_2), \dots$ we get a downset $\mathcal{D}$ and by (1.21) $|\mathcal{D}| = |\mathcal{C}|$ and $\mathcal{D} = \text{Ł}(\mathcal{D}) \subset \text{Ł}(\mathcal{C})$, which proves Theorem 3.

We are going now for equality characterization in (1.18). We already now equality for downsets.

Symmetrically, if $\mathcal{C}$ is an upset, then $\text{Ł}(\mathcal{C}) = \bar{\mathcal{C}} = \{[n] \smallsetminus C : C \in \mathcal{C}\}$ and again there is equality.

Defining push–up operation U analogously to the push–down operation T, then by symmetry

$$|U(\mathcal{C}, x)| = |\mathcal{C}| \text{ and } \text{Ł}(U(\mathcal{C}, x)) \subset \text{Ł}(\mathcal{C}). \tag{1.22}$$

We say now that a set $\mathcal{A} \subset 2^{[n]}$ is accessible for $\mathcal{C} \subset 2^{[n]}$, if starting from $\mathcal{C}$ one can obtain $\mathcal{A}$ by consecutively applying finitely many push–down and push–up operations. (For example $T(\mathcal{C}, x_1), \mathcal{U}(T(\mathcal{C}, x_1), x_2), T(U(T(\mathcal{C}, x_1), x_2), x_3), \dots$.

By (1.21), (1.22) $|\mathcal{A}| = |\mathcal{C}|$, if $\mathcal{A}$ is an accessible set for $\mathcal{C}$.

Theorem 4. *Let $\mathcal{D}_1, \dots, \mathcal{D}_m \subset 2^{[n]}$ be the downsets accessible for a set $\mathcal{C} \subset 2^{[n]}$, then*

$$\bigcup_{i=1}^{m} \mathcal{D}_i = \text{Ł}(\mathcal{C}). \tag{1.23}$$

In particular $|\text{Ł}(\mathcal{C})| = |\mathcal{C}|$ if and only if there exists a unique downset accessible for $\mathcal{C}$.

This result holds, because by (1.21), (1.22) any downset $\mathcal{D}$ accessible for $\mathcal{C}$ satisfies $\mathcal{D} \subset \text{Ł}(\mathcal{C})$ and thus $\bigcup_{i=1}^{m} \mathcal{D}_i \subset \text{Ł}(\mathcal{C})$. Further, for any $S \in \text{Ł}(\mathcal{C})$ we get from

$\mathcal{C}$ by down pushing in all $x \in \bar{S}$ downset $\mathcal{D}_S$ with $S \in \mathcal{D}_S$ and thus equality in (1.23).

Moreover, if there is exactly one $\mathcal{D}_i$, then $|\mathrm{L}(\mathcal{C})| = |\mathcal{D}_i| = |\mathcal{C}|$ and, conversely the equation $|\mathrm{L}(\mathcal{C})| = |\mathcal{C}|$ implies that $m = 1$, because the union of two or more sets of cardinality $|\mathcal{C}|$ each, would imply $\left| \bigcup_{i=1}^{m} \mathcal{D}_i \right| > |\mathcal{C}|$.

However, in order to understand the structure of $\mathcal{C}'$ with equality in (1.18) it is important to notice that **accessibility is not commutative**: if $\mathcal{D}$ is accessable for $\mathcal{C}$, then $\mathcal{C}$ needs not be accessable for $\mathcal{D}$.

Example 3: Let $n = 3$, $\mathcal{C} = \{\phi, \{1\}, \{2, 3\}, \{3\}\}$, then $\mathrm{L}(\mathcal{C}) = \{\phi, \{1\}, \{2\}, \{3\}\}$ and $|\mathrm{L}(\mathcal{C})| = |\mathcal{C}| = 4$. However, $\mathcal{C}$ is not accessible for downsets and upsets as can be checked.

2 Main Auxiliary Old and New Results for the Proof of Theorem 2

We make essentially use of Harper's Edge Isoperimetric Theorem ([1], correct proofs in [2] – [3]). For $C \subset \mathcal{X}^n$ define

$$\emptyset(C) = \big\{ (x^n, y^n) : x^n \in C, y^n \in \mathcal{X}^n \smallsetminus C, d(x^n, y^n) = 1 \big\}. \tag{2.1}$$

For a number M define

$$\omega(M) = \min_{C \subset \mathcal{X}^n, |C| = M} |\emptyset(C)|. \tag{2.2}$$

Theorem (Edge Isoperimetry). *The minimum in (2.2) is assumed for a generalized cylinder.*

We recall the definition of a generalized cylinder.

Every positive integer M can uniquely be written in a binary expansion

$$M = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_s}, n_1 > n_2 > \cdots > n_s \geq 0. \tag{2.3}$$

Let us use the picture in (2.4) for the set $\mathcal{X}^m$, then we can present

$$Z(M) = \mathcal{X}^{n_1} \times \{0\}^{n - n_1}$$
$$\dot{\cup}\, \mathcal{X}^{n_2} \times \{0\}^{n - n_2} \times \{1\} \times \{0\}^{n - n_1 - 1}$$
$$\vdots$$

as union of the disjoint sets

$$
\begin{array}{|l|l}
\hline
n_1 & \quad\quad 000\ldots 0 \\
\hline
n_2 & \quad 000100\ldots 0 \\
\hline
n_3 & \; 00100100\;\ldots 0 \\
\hline
n_4 & 0010100100\;\;\ldots 0 \\
\hline
\end{array}
\qquad (2.4)
$$

$$\vdots$$

Instead of minimizing the number $\mathrm{out}(C) = |\varnothing(C)|$ of outgoing "edges" we can equivalently maximize the number $\mathrm{int}(C) = |I(C)|$ of internal "edges", because all vertices have degree n and thus

$$
\mathrm{out}(C) + \mathrm{int}(C) = n|C|.
$$

We refer then to the dual form of this optimisation problem.

Define

$$
i(M) = \max_{C:|C|=M} \mathrm{int}(C). \qquad (2.5)
$$

For the proof of our Uniqueness Theorem below we need

Lemma 1. *For $M < 2^k$ necessarily*

$$
i(M) < k \cdot \frac{M}{2}.
$$

Proof: We can calculate $i(M)$ from Theorem H_1 using the representation (2.4). Indeed

$$
\begin{aligned}
i(M) = {} & n_1\, 2^{n_1-1} + n_2\, 2^{n_2-1} + n_3\, 2^{n_3-1} + \cdots + n_s\, 2^{n_s-1} \\
& + 2^{n_2} + 2 \cdot 2^{n_3} + \cdots + (n_s - 1)2^{n_s}. \qquad (2.6)
\end{aligned}
$$

For $M < 2^k$ necessarily

$$
k \geq n_1 + 1 \geq n_2 + 2 \geq n_3 + 3 \geq \ldots \qquad\qquad . \qquad (2.7)
$$

We have to show that $k\, M > 2\, i(M)$ or that

$$
\begin{aligned}
k(2^{n_1} + 2^{n_2} + \cdots + 2^{n_s}) > {} & n_1\, 2^{n_1} + n_2\, 2^{n_2} + \cdots + n_s\, 2^{n_s} \\
& + 1 \cdot 2 \cdot 2^{n_2} + 2 \cdot 2 \cdot 2^{n_3} + 3 \cdot 2 \cdot 2^{n_4} + \ldots
\end{aligned}
$$

or that

$$
(k-n_1) \cdot 2^{n_1} + (k-n_2)2^{n_2} + \cdots + (k-n_s)2^{n_s} > 1 \cdot (2 \cdot 2^{n_2}) + 2(2 \cdot 2^{n_3}) + \cdots +. \qquad (2.8)
$$

Now, by (2.7) $(k - n_i) \geq i$ and the RHS in (2.8) does not exceed the LHS even if we ignore the term $(k - n_s)2^{n_s}$, which is positive. Thus (2.8) holds and the Lemma is proved.

Uniqueness Theorem. *Generalized cylinders are up to permutations $\pi \in \Sigma_n$ and additions with x^n in $GF(2)^n$ the only solutions in Theorem H_2 (at least for $M = 2^k$).*

Proof for the case $M = 2^k$: Let $C \subset \mathcal{X}^n$ be optimal. Then

$$
\mathrm{int}(C) = i(M) = k \cdot 2^{k-1}. \qquad (2.9)
$$

Consider

$$C_1 = \big\{(x_2,\ldots,x_n):(1,x_2,\ldots,x_n)\in C\big\},\, C_0 = \big\{(x_2,\ldots,x_n):(0,x_2,\ldots,x_n)\in C\big\}. \tag{2.10}$$

We proceed by induction on n.

We have to show that either $C_1 = \varnothing$ or $C_0 = \varnothing$ or that $C_1 = C_0$.

Case $|C_1| = |C_0|$:
Consider

$$\begin{aligned}
\mathrm{int}(C) &= \mathrm{int}(C_1) + \mathrm{int}(C_0) + |C_1 \cap C_0| \le 2i(2^{k-1}) + |C_1 \cap C_0| \\
&\le 2(k-1)2^{k-2} + |C_1 \cap C_0| = (k-1)2^{k-1} + |C_1 \cap C_0| = k \cdot 2^{k-1}.
\end{aligned} \tag{2.11}$$

Since $\mathrm{int}(C) = k \cdot 2^{k-1}$, necessarily $|C_1 \cap C_0| = 2^{k-1}$ and thus $C_1 = C_0$.

Case $|C_1| > |C_0|$:
Here $|C_1| = 2^{k-1} + 2^{\ell_2} + \ldots$, $\ell_2 > 0$, and $|C_0| = 2^k - |C_1|$.

Subcase $2^k > |C_1| = 2^{k-1} + 2^{k-2} + \ldots$:
Here $|C_0| \le 2^{k-3} + \cdots < 2^{k-2}$.

By Lemma 1 $\mathrm{int}(C_1) < k\frac{|C_1|}{2}$ and $\mathrm{int}(C_0) \le (k-2)\frac{|C_0|}{2}$.

Therefore by (2.11)

$$\mathrm{int}(C) < k\frac{|C_1|}{2} + (k-2)\frac{|C_0|}{2} + |C_0| = k\frac{|C|}{2} = k \cdot 2^{k-1},$$

which contradicts (2.9).

Subcase $2^k > |C_1| = 2^{k-1} + 0 + \ldots$:
Here $|C_0| = 2^{k-2} + \cdots < 2^{k-1}$.

The estimate of $\mathrm{int}(C_1)$ is more tricky. We use the representation (2.4) for C_1 and obtain

$$\mathrm{int}(C_1) \le (k-1)2^{k-2} + (|C_1| - 2^{k-1}) + i(|C_1| - 2^{k-1}). \tag{2.12}$$

Since by Lemma 1 $i(|C_1| - 2^{k-1}) < (k-2)\frac{|C_1| - 2^{k-1}}{2}$ and $\mathrm{int}(C_0) < (k-1)\frac{|C_0|}{2}$, we have

$$\mathrm{int}(C) < (k-1)\left(2^{k-2} + \frac{|C_0|}{2}\right) + k\frac{|C_1| - 2^{k-1}}{2} = k\frac{|C_0| + |C_1|}{2} - \left[\frac{|C_0|}{2} - (k-1)2^{k-2} + k\frac{2^{k-1}}{2}\right]$$

in contradiction to $\mathrm{int}(C) = k \cdot 2^{k-1}$, because the term in brackets is positive.

3   Further Auxiliary Results

We use again the dual form and apply it to $C = A \cup B$.

Lemma 2. *For an admissible k–pair (A, B) we have*

$$|A \cup B| \geq 2^k. \tag{3.1}$$

Proof: Let $e(A, B)$ be the number of edges between A and B. Then by our assumptions

$$\mathrm{int}(C) \geq e(A, B) \geq k|B|.$$

Since also $|B| \geq |A|$, therefore also

$$\mathrm{int}(C) \geq k\frac{|C|}{2}. \tag{3.2}$$

We assume now that (3.1) does not hold and derive a contradiction. If now $M = |C| < 2^k$, then by (3.2)

$$i(M) \geq \mathrm{int}(C) \geq k\frac{M}{2}. \tag{3.3}$$

Next we use this result to derive a lower bound on $|A|$.

Lemma 3. *For an admissible k–pair (A, B) necessarily*

$$|A| \geq 2^{k-1}.$$

Proof: For an admissible k–pair (A, B) label the elements in B as $b_1 \ldots b_r$ such that

$$|A \cap S(b_i)| \leq |A \cap S(b_{i+1})| \text{ for } i = 1, \ldots, r - 1$$

and define for $j = |B| - |A|$

$$B' = B - \{b_1, \ldots, b_j\}.$$

Notice that (A, B') is an admissible k–pair with $|A| = |B'|$ and thus $|A| \geq \frac{|A \cup B'|}{2}$. Since by Lemma 2 $|A \cup B'| \geq 2^k$, the result follows.

Lemma 4. *For an optimal k–pair (A, B) necessarily*
  (a) $|A| = |B| = 2^{k-1}$
  (b) $A \cap B = \varnothing$.

Proof: We know from Lemma 3 that (Z_+^k, Z_-^k) is a minimal k–pair and thus $|A| = 2^{k-1}$.
  Suppose that $|B| = 2^{k-1} + b$, $b > 0$.
  Then by the procedure described in the proof of Lemma 2 we can get a minimal k–pair (A, B') with $|B'| = 2^{k-1} + 1$.

By assumption

$$\text{int}(A \cup B') \geq \sum_{b \in B'} |A \cap S(b)| \geq k|B'| \geq k \cdot 2^{k-1} + k. \tag{3.4}$$

However, since $|A \cup B'| \leq 2^k + 1$ we have

$$\text{int}(A \cup B') \leq i(2^k + 1) = k \cdot 2^{k-1} + 1. \tag{3.5}$$

Now (3.5) contradicts (3.4) for $k \geq 2$. (b) follows with Lemma 2.

Remark: For $k = 1$ $\left(\{00\ldots0\}, \{x^n : \sum_{t=1}^n x_t = 1\}\right)$ is an optimal 1–pair and (a) does not hold!

4   Proof of Theorem 2

We know already that for an optimal k–pair (A, B) necessarily $A \cap B = \varnothing$, $|A| = |B| = 2^{k-1}$. Since also $\text{int}(A \cup B) \geq k|B| = k2^{k-1}$ and by Theorem H_1 $\text{int}(A \cup B) \leq i(2^k) = k \cdot 2^{k-1}$, we conclude that

$$\text{int}(A \mathbin{\dot{\cup}} B) = k \cdot 2^{k-1}. \tag{4.1}$$

By the Uniqueness Theorem $A \cup B$ is a cylinder Z and w.l.o.g. $Z = \boxed{k}\,00\ldots0$. Since every element in Z has degree k and $|B|k = 2^{k-1} \cdot k$, necessarily

$$|A \cap S(b)| = k \text{ for all } b \in B. \tag{4.2}$$

Now (4.1) and (4.2) imply

$$e(A, B) = \text{int}(A \cup B), \tag{4.3}$$

that is, all edges are between the sets (or no edges are in A or in B). Clearly

$$|B \cap S(a)| = k \text{ for all } a \in A. \tag{4.4}$$

Finally, we can assume w.l.o.g. that $00\ldots0 \in B$. Then no singleton can be in B and by (4.2) the first k of the n singletons must be in A.

By the degree condition (4.4) then all doubletons with 1's in $\{1, 2, \ldots, k\}$ must be in B, next all tripletons with 1's in $\{1, 2, \ldots, k\}$ must be in A, etc. This completes the proof.

Problem 4: How do the results generalize from $\mathcal{X} = \{0, 1\}$ to $\mathcal{X} = \{0, \ldots, \alpha-1\}$?

5   Further Observations

On large boundaries of intensity k.
    We adapt the *convention*: $x = x^n$.
    Recall the definition of $\tilde{V}_k$ in (1.13). For $1 \leq k \leq n$ and $0 \leq N \leq 2^n$ we define now

$$\gamma(n, N, k) = \max_{|V| = N} |\tilde{V}_k|. \tag{5.1}$$

Fact I: γ is not decreasing in N. Unlike Theorem 1 it is here very difficult to obtain exact results. We discuss therefore some special cases of the function γ.

Case $k = 1$: For not too large N it is clearly optimal to choose V as a 1–error correcting code, that is,

$$\min_{x,y \in V} d(x, y) \geq 3. \tag{5.2}$$

Here $|\tilde{V}| = n|V|$ and this is optimal.

Case $k \geq 2$: The situation is now quite different, because the points in V should be not too far apart

Fact II: The points in $A(V) = \{x \in V : d(V \smallsetminus \{x\}, x) \geq 2\}$ don't contribute to $\tilde{V}_k$ for $k \geq 2$. We can therefore assume that $A(V) = \varnothing$.

Fact III: Let us associate with V the graph $G_2 = G_2(V) = (V, \mathcal{E})$, where

$$\mathcal{E} = \big\{\{x, y\} : x, y \in V \text{ and } d(x, y) = 2\big\}.$$

We can assume that $G_2(V)$ is *connected*, because $x \in \mathcal{X}$ can have distance 1 only with vertices in the same connected component.

Case $k = 2$: By induction on n one gets

Lemma 5. *If $G_2(V)$ is connected and $|V| \geq 2$, then $|V| \leq |\tilde{V}_2|$.*
    We derive now a lower bound on $\alpha(n, N, k)$ for smaller k.

Lemma 6. *For $k \leq \log N + 1$*

$$\alpha(n, N, k) = \Omega\left(\frac{N \cdot \log N}{k - 1}\right).$$

Proof: Being concerned only about the **order of growth** we make the simplifying assumption

$$k - 1 \,|\, \log N.$$

Recall the definition of $\mathcal{X}_+^k$ in Section 3 and choose $C = \frac{\log N}{k-1}$ of its copies $\mathcal{X}_{+\ell}^k$, $1 \leq \ell \leq L$, and define $W^k = \mathcal{X}_{+1}^k \times \cdots \times \mathcal{X}_{+L}^k \times \{0\} \times \cdots \times \{0\} \subset \mathcal{X}^n$, where the factor $\{0\}$ occurs exactly $n - L \cdot k$ times.
    Clearly, $|W^k| = 2^{(k-1)L} = N$ and for $V = W^k$ we have

$$\tilde{V}_k = \{x \oplus e_j : x \in W^k, 1 \leq j \leq kL\},$$

where e_j has a 1 in the j–th position and 0 otherwise.
    Hence $|\tilde{V}_k| = |W^k|\frac{kL}{k} = NL$, because $\mathcal{X}_{+\ell}^k + e_i = \mathcal{X}_{+\ell}^k + \ell$, for $k(\ell - 1) + 1 \leq i, j \leq k\ell$.

On a dual form of the vertex isoperimetric theorem in the Hamming space $(\mathcal{X}^n, d)$.

(A, B) with $A, B \subset \mathcal{X}^n$ is an (n, d)–pair, if

$$d_H(a, b) \geq d \text{ for } a \in A, b \in B.$$

It was shown in [AK] that

$$\max\{|A||B| : (A, B) \text{ is } (n, d)\text{–pair}\}$$

is assumed for $(A, B) = \{a_t^n \mathcal{X}^n : w(a^n) \leq \ell\}$, $\{b^n \in \mathcal{X}^n : w(b^n) \geq d + \ell\}$ with a suitable ℓ.

Here we show that $\ell = \lceil \frac{n-d}{2} \rceil$.

Lemma 7. *For every $n \in \mathbb{N}$, $0 \leq d \leq n$, and $0 \leq \ell \leq n$ the values $f_{n,d}(\ell) =$*

$$\left(\sum_{i=0}^{\ell} \binom{n}{i} \right) \left(\sum_{i=d+\ell}^{n} \binom{n}{i} \right) \text{ satisfy}$$

$$f_{n,d}\left(\left\lceil \frac{n-d}{2} \right\rceil \right) \geq f_{n,d}(\ell).$$

Proof: By symmetry it suffices to show that $f_{n,d}(\ell) \leq f_{n,d}(\ell+1)$, if $n \geq 2\ell+1+d$, that is,

$$\left(\sum_{i=0}^{\ell} \binom{n}{i} \right) \left(\sum_{i=d+\ell}^{n} \binom{n}{i} \right) \leq \left(\sum_{i=0}^{\ell} \binom{n}{i} \right) \left(\sum_{i=d+\ell+1}^{n} \binom{n}{i} \right)$$

iff

$$\left(\sum_{i=0}^{\ell} \binom{n}{i} \right) \binom{n}{d+\ell} \leq \binom{n}{\ell+1} \left(\sum_{i=d+\ell+1}^{n} \binom{n}{i} \right). \tag{5.3}$$

We prove inequality (5.3), by induction on n for all ℓ, d with $n \geq 2\ell + 1 + d$.

For $n = 2$, that is, $\ell = 0$, (5.3) obviously holds. Therefore we assume (5.3) to be true for $n' < n$. We consider first the cases $\ell = 0$, and $n = 2\ell + 1 + d$.

a) $d = 1$:

$$\left(\sum_{i=0}^{\ell} \binom{n}{i} \right) \binom{n}{\ell+1} \leq \binom{n}{\ell+1} \left(\sum_{i=\ell+2}^{n} \binom{n}{i} \right).$$

b) $\ell = 0$:

$$\binom{n}{0}\binom{n}{d} = \binom{n}{d} \leq \binom{n}{1} \left(\sum_{i=d+1}^{n} \binom{n}{i} \right) = n \binom{n}{d+1} + \sum_{i=d+2}^{n} \binom{n}{i}.$$

Since $\binom{n}{d} \leq n\binom{n}{d+1}$ (∗) holds for $\ell = 0$.

c) $n = 2\ell + 1 + d$ implies

$$\left(\sum_{i=0}^{\ell}\binom{n}{i}\right)\binom{n}{d+\ell} = \left(\sum_{i=0}^{\ell}\binom{n}{i}\right)\binom{n}{\ell+1} = \left(\sum_{i=d+\ell+1}^{n}\binom{n}{i}\right)\binom{n}{\ell+1}.$$

Let now $\ell \geq 1$, $d \geq 2$, $n \geq 2\ell + d + 2$:

$$\left(\sum_{i=0}^{\ell}\binom{n}{i}\right)\binom{n}{d+\ell} = \left(\sum_{i=0}^{\ell}\binom{n-1}{i} + \sum_{i=0}^{\ell-1}\binom{n-1}{i}\right)\left(\binom{n-1}{d+\ell} + \binom{n-1}{d+\ell-1}\right)$$

$$= \left(\sum_{i=0}^{\ell}\binom{n-1}{i}\right)\binom{n-1}{d+\ell} + \left(\sum_{i=0}^{\ell}\binom{n-1}{i}\right)\binom{n-1}{d+\ell-1} + \left(\sum_{i=0}^{\ell-1}\binom{n-1}{i}\right)\binom{n-1}{d+\ell}$$

$$+ \left(\sum_{i=0}^{\ell-1}\binom{n-1}{i}\right)\binom{n-1}{d+\ell-1}$$

$$\leq \binom{n-1}{\ell+1}\left(\sum_{i=d+\ell+1}^{n-1}\binom{n-1}{i}\right) + \binom{n-1}{\ell+1}\left(\sum_{i=d+\ell}^{n-1}\binom{n-1}{i}\right) + \binom{n-1}{\ell}\left(\sum_{i=d+\ell+1}^{n-1}\binom{n-1}{i}\right)$$

$$+ \binom{n-1}{\ell}\sum_{i=d+\ell}^{n-1}\binom{n-1}{i}$$

$$= \left(\binom{n-1}{\ell+1} + \binom{n-1}{\ell}\right)\left(\sum_{i=d+\ell+1}^{n-1}\binom{n-1}{i} + \sum_{i=d+\ell}^{n-1}\binom{n-1}{i}\right) = \binom{n}{\ell+1}\left(\sum_{i=d+\ell+1}^{n}\binom{n-1}{i}\right).$$

Remark: The result must be known and also have a simpler proof!

6   Concluding Conjectures

On k–pairs

We consider

$$\alpha(n, N, k) = \max\{|B| : A, B \subset \mathcal{X}^n, |A| \leq N, (A, B) \text{ is } k\text{–pair}\}.$$

Conjectures

The following constructions give the "asymptotic" value of $\alpha(n, N, k)$:

a.) If $N = \binom{m}{k-1}$ and $2(k-1) < m \leq n$, then $A = \binom{[m]}{k-1} \times \binom{[n-m]}{0}$, $B = \binom{[m]}{k} \times \binom{[n-m]}{0}$

$\alpha(n, N, k) = N \cdot \frac{m-k+1}{k}$ $m = 2k = n$?

b.) If $N = \binom{k}{j} + \binom{k}{j-2}$ with $2 \leq j \leq \frac{k}{2}$, then $\alpha(n, N, k) \sim N\frac{j}{k-j+1}$.

Use $B = \binom{[k]}{j-1} \times \binom{[n-k]}{0}$, $A = \left(\binom{[k]}{j} \cup \binom{[k]}{j-2}\right) \times \binom{[n-k]}{0}$.

On edge isoperimetry

For $A \subset \mathcal{X}^n$ define $E_A = \big\{(x, y) : x, y \in A, \ x \oplus y = e_i \text{ for some } i\big\}$, where $e_i = (0, 0, \ldots, 1, 0, \ldots, 0)$, with "1" in component i. Then

$$f_2(N) = \max_{\substack{A \subset \mathcal{X}^n \\ |A| = N}} |E_A| f_2(2^k) = k \cdot 2^k \text{ by Harper and } f_2(N) \le N \log N. \quad (6.1)$$

Define now $E_r(A) = \left\{(x_1, x_2, \ldots, x_r) : x_i \in A, \ \bigoplus_{j=1}^{r} x_j = e_i \text{ for some } i\right\}$ and

$$f_r(N) = \max_{\substack{A \subset \mathcal{X}^n \\ |A| = N}} |E_r(A)|.$$

We can write $E_r(A) = \overset{\cdot}{\underset{x \in A}{\bigcup}} B_x$, where

$$B_x = \left\{(x, x_2, \ldots, x_r) : \bigoplus_{j=2}^{r} x_j = e_i \oplus x\right\} \text{ and}$$

$$|B_x| = \left|\left\{(x_2, \ldots, x_r) : \bigoplus_{j=2}^{r} x_j = e_i\right\}\right| \le f_{r-1}.$$

Consequently $f_r(2^k) = (2^k)^{r-1} \cdot k$ (Harper's cylinder).

Conjecture

$$f_r(N) \le N f_{r-1}(N) \le \cdots \le N^{r-2} f_2(N) \le N^{r-1} \log N.$$

References

1. L.H. Harper, Optimal assignment of numbers to vertices, J. Soc. Industr. Appl. Math., 12, 1, 385–393, 1964.
2. H.H. Lindsey, Assignment of numbers to vertices, Amer. Math. Monthly 71, 508–516, 1964.
3. A.J. Bernstein, Maximal connected arrays on the n–cube, SIAM J. Appl. Math. 15, 6, 1485–1489, 1967.
4. L.H. Harper, Optimal numberings and isoperimetric problems on graphs, J. Combin. Theory 1, 385–393, 1966.
5. R. Ahlswede and Z. Zhang, An identity in combinatorial extremal theory, Adv. in Math., Vol. 80, No. 2, 137–151, 1990.
6. R. Ahlswede and Z. Zhang, On multi–user write–efficient memories, IEEE Trans. Inform. Theory, Vol. 40, No. 3, 674–686, 1994.
7. R. Ahlswede and N. Cai, Models of multi–user write–efficient memories and general diametric theorems, Preprint 93–019, SFB 343 "Diskrete Strukturen in der Mathematik", Universität Bielefeld, Information and Computation, Vol. 135, No. 1, 37–67, 1997.

8. R. Ahlswede and G. Katona, Contributions to the geometry of Hamming spaces, Discrete Mathematics 17, 1–22, 1977.

9. R. Ahlswede, Advances on extremal problems in number theory and combinatorics, European Congress of Mathematics, Barcelona 2000, Vol. I, 147–175, Carles Casacuberta, Rosa Maria Miró–Roig, Joan Verdera, Sebastiá Xambó–Descamps, edit., Progress in Mathematics, Vol. 201, Birkhäuser Verlag, Basel–Boston–Berlin, 2001.

10. J. Lawrence, Lopsided sets and orthant–intersection of convex sets, Pacific J. Math. 104, 155–173, 1983.

11. H.J. Bandelt, V. Chepoi, A. Dress, and J. Koolen, Theory of lopsided set systems, Preprint.