Discrete Mathematics 17 (1977) 1–22.

CONTRIBUTIONS TO THE GEOMETRY OF HAMMING SPACES*

R. AHLSWEDE

Fakultät für Mathematik, 4800 Bielefeld, W. Germany

G.O.H. KATONA

Mathematical Institute of the Hungarian Academy of Sciences, 1053 Budapest, Hungary

Received 16 December 1975

Let $1 \leq N \leq 2^n$ and let $\mathfrak{A}, \mathfrak{B}$ denote families of subsets of $\{1, \ldots, n\}$. The following results are proved:

Theorem 2.3. *$(\mathfrak{A}, \mathfrak{B})$ is a d-pair, $1 \leq d \leq n$, if $|A \Delta B| \leq d$ for all $A \in \mathfrak{A}$, $B \in \mathfrak{B}$. Then $\max\{|\mathfrak{B}| : (\mathfrak{A}, \mathfrak{B})$ is d-pair and $|\mathfrak{A}| = N\}$ is assumed if $\mathfrak{A}$ is a "quasi-sphere".*

Theorem 3.1. *$\min_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_{A,B\in\mathfrak{A}} |A \cap B|$ is assumed for a (pseudo)-sphere characterized by the property that*

$$\big||\{A : A \in \mathfrak{A}, x \in A\}| - |\{A : A \in \mathfrak{A}, y \in A\}|\big| \leq 1$$

for all $x, y \in \{1, 2, \ldots, n\}$.

Denote by $K_i = K_i(\mathfrak{A})$ $(i = 0, 1, \ldots, n)$ the number of i-element members of an order ideal $\mathfrak{A}$.

Theorem 4.2. *$\min_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_i K_i W_i$ is assumed*

(a) *in case $W_0 \leq W_1 \leq \cdots \leq W_n$ if $\mathfrak{A}$ is a quasi-sphere,*

(b) *in case $W_0 \geq W_1 \geq \cdots \geq W_n$ if $\mathfrak{A}$ is a quasi-cylinder.*

Theorem 4.5. *$\min_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_i K_i W_i$ is assumed*

(a) *in case $W_0 \leq W_1 \leq \cdots W_{M+1} \geq \cdots \geq W_n$ by a union of a quasi-cylinder and a quasi-sphere,*

(b) *in case $W_0 \geq W_1 \geq \cdots \geq W_M \leq W_{M+1} \leq \cdots \leq W_n$ by an intersection of a quasi-cylinder and a quasi-sphere.*

1. Introduction

Let $H^n = \prod_1^n \{0, 1\}$ be the set of (0–1)-sequences of length n[1] and let d denote the Hamming metric in H^n, that is, for any two elements $x^n = (x_1, \ldots, x_n)$, $y^n = (y_1, \ldots, y_n) \in H^n$,

* The Research of the first author was supported by the Deutsche Forschungsgemeinschaft. The work of the second author was done while the author visited the Institute of Mathematical Statistics, University of Göttingen.

[1] There is a natural correspondence between those (0–1)-sequences and the subsets of the set $\{1, \ldots, n\}$. We shall use both — the sequence and the subsets — terminology in this paper. Sets of cardinality k are also called k-tuples.

$$d(x^n, y^n) = |\{t : x_t \neq y_t,\ 1 \leq t \leq n\}|. \tag{1.1}$$

This is a very canonical metric for (0–1)-sequences and was used by Hamming for his investigations in the theory of error correcting codes [14]. We refer to (H^n, d) shortly as Hamming space. Whereas the theory of error correcting codes has used since its early days algebraic and also combinatorial techniques, this used to be not the case for the Shannon theory of communication, which heavily depends on probabilistic methods. However recent investigations in multi-user communication ([1, 2]) led to new combinatorial extremal problems in (H^n, d) and also more general spaces, which we shall not consider here. Related problems occurred in the study of random graphs [26]. We state one important result of [26] (in the slightly improved form of [2]), which seems to be of general interest to probability theory.

Let $P^n = P \times \cdots \times P$ be a product distribution on $X^n = \prod_1^n X$, $X = \{1, \ldots, a\}$, and define for $B \subset X^n$:

$$\Gamma^k B = \{x^n : x^n \in X^n,\ d(x^n, y^n) \leq k \text{ for some } y^n \in B\} \tag{1.2}$$

for $k = 1, 2, \ldots$.

Furthermore, define the "inner" surface δB of B by

$$\delta B = B \cap \Gamma^1 \bar{B}, \tag{1.3}$$

where here and elsewhere $\bar{A}$ denotes the complement of a set A. Then,

$$P^n(\delta B) \geq c \cdot n^{-\frac{1}{2}} f(P^n(B)), \tag{1.4}$$

where $f(s) = \phi(\Phi^{-1}(s))$, $\phi(t) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}t^2}$, $\Phi(t) = \int_{-\infty}^{t} \phi(u)du$ and where c is a constant depending only on P.

As a consequence of (1.4) one obtains (see [2]) for all $B \subset X^n$,

$$P^n(\Gamma^k B) \geq \Phi[\Phi^{-1}(P^n(B)) + n^{-\frac{1}{2}}(k-1)c], \qquad k = 1, 2, \ldots. \tag{1.5}$$

Those results are exact up to a multiplicative constant and imply that in case $X = \{0, 1\}$ and $P(0) = P(1) = \frac{1}{2}$ among all the subsets of H^n with given cardinality the Hamming spheres have smallest (within an accuracy given by this constant) surface. (The surface of a set B is $\Gamma B - B$.) The same result holds for the k-surface $\Gamma B^k - B$. This result suggests that the sphere is the exact solution to the problem and indeed this was proved to be true in [16]. More specifically the following was proved.

Theorem 1.1. *If $A \subset H^n$, $|A| = N < 2^n$, then*

$$|\Gamma^d(A)| \geq G_d(n, N) \quad \text{for } d = 1, 2, \ldots, \tag{1.6}$$

where

$$G_d(n, N) = \binom{n}{n} + \cdots + \binom{n}{k+1} + \cdots + \binom{n}{k-d+1} + \binom{a_k}{k-d} + \cdots + \binom{a_t}{t-d},$$

if N is (uniquely) represented as

$$N = \binom{n}{n} + \cdots + \binom{n}{k+1} + \binom{a_k}{k} + \cdots + \binom{a_t}{t}$$

for some k; $t - 1 \leqslant k \leqslant n$; *and*

$$n > a_k > a_{k-1} > \cdots > a_t \geqslant t \geqslant 1.$$

Equality holds in (1.6) if A consists of all l-tuples; $0 \leqslant l \leqslant k+1$; and $\binom{a_k}{k} + \cdots + \binom{a_t}{t}$ $(k+2)$-tuples chosen in lexicographical order. We refer to such a set as a quasi-sphere and if $\binom{a_k}{k} + \cdots + \binom{a_t}{t} = 0$ as a sphere of Hamming radius $k+1$ and center $\mathbf{0} = (0, \ldots, 0)$. H^n can be viewed as vector space over the field GF(2) and the metric d is invariant under translation by a vector. Therefore the above statements apply to spheres or quasi-spheres with any center. The special case $d = 1$ has a striking interpretation and simply means that given the cardinality ("volume") the sphere has minimal cardinality of the surface. This phenomenon is known as isoperimetric property for euclidean [30] and also non-euclidean geometries (see [28, 29]). Since (H^n, d) is isomorphic to the family of subsets of an n-set endowed with the symmetric difference as distance function every result about (H^n, d) has directly a set theoretic or combinatorial interpretation. The classical isoperimetric property has been studied in great detail and many consequences have been derived. Those geometric results can now serve as guides for finding analogous combinatorial results or at least to derive some of the known results by a unified approach. Combinatorics has always earned the criticism of lacking general theories and the present attempt may help to carry some general principles into the area of extremal problems. As far as our actual results go this is just a beginning.

In Section 2 we show that earlier results (see [17, 20]) can be derived from the isoperimetric property and can be stated as: for given "volume" the Hamming sphere has minimal diameter.

The "Spiegeltheorem" of Schmidt [28], a dual form of the isoperimetry theorem has a simple analogue in H^n and leads to a combinatorial result, which was previously unknown (Theorem 2.3, Section 2).

In Section 3 we investigate sum type extremal problems, which add a new dimension to extremal problems considered so far in the literature (see [10, 18, 12]). Theorem 3.1 gives a new characterisation of the sphere. The concept of an order ideal is very basic for many combinatorial problems (see [12]). They are defined for partially ordered sets and naturally extend the notion of a simplicial complex to which they specialize in H^n. They provide the answer to many extremal problems and deserve a study on their own. In Section 4, Theorems 4.2 and 4.5, we give geometric characterisations of order ideals, which are optimal under certain weight assignments to the "levels" of the ideals. In Section 5 an application to random graphs is given.

2. New and old combinatorial results as consequences of the isoperimetric property in Hamming spaces

We begin with some general remarks about the isoperimetric property.

For a set $A \subset H^n$ one can write $\Gamma^r(A)$ (defined in (1.2)) also as

$$\Gamma^r(A) = \bigcup_{x^n \in A} S_r(x^n), \tag{2.1}$$

where

$$S_r(x^n) = \{y^n : y^n \in H^n,\ d(x^n, y^n) \leqslant r\}.$$

Since

$$S_r(x^n) = x^n + S_r(\mathbf{0}) \tag{2.2}$$

with the addition understood in the vector space H^n, one can also express $\Gamma^r(A)$ as the Minkowski sum (that is $B + C = \{b + c : b \in B, c \in C\}$)

$$\Gamma^r(A) = A + S_r(\mathbf{0}). \tag{2.3}$$

The isoperimetric property then means that $\min_{A:|A|=N} |A + S_r(\mathbf{0})|$ is assumed for a quasi-sphere. In this formulation one easily recognizes a similarity to the Brunn-Minkowski inequality (B.M.I.) for the euclidean space E^n [5, 27], especially if one uses a formulation due to Schmidt [29].

For any Lebesgue measurable set $A \subset E^n$ he defines the radius $v(A)$ as the radius of a sphere, whose volume equals the volume of A. The B.M.I. then takes the form

$$v(A + B) \geqslant v(A) + v(B) \tag{2.4}$$

with equality if and only if A and B are spheres up to null sets. In [28] a closely related result was obtained, called the Spiegeltheorem,

$$v(A/B) \leqslant v(A) - v(B) \tag{2.5}$$

with equality if and only if A and B are spheres up to null sets. Here $A/B = \bigcap_{b \in B}\{-b + A\}$ is the Minkowski difference.

In [28] and [29] those inqualities were extended to all non-euclidean geometries.

The structure of the space H^n is quite different, however. The most apparent and most basic differences are:

(a) H^n is discrete (even finite) — E^n is nondiscrete.

(b) Complements of spheres are spheres in H^n — this property obviously does not hold in E^n.

(c) H^n has subgroups, whereas E^n has no subgroup of positive finite measure.

An immediate consequence of (c) is that the B.M.I. cannot hold for H^n in its full generality.

However, there are important special cases, which are obtained by assuming that one of the two sets is a sphere. For $A \subset E^n$ define

$$\Gamma^r A = \{x : x \in E^n,\ \rho(x, A) \leqslant r\},$$
$$\Gamma_r A = \{x : x \in E^n,\ \rho(x, a) \leqslant r \quad \forall a \in A\},$$
$$\Gamma_{-r} A = \{x : x \in E^n,\ \rho(x, \bar{A}) \geqslant r\}, \tag{2.6}$$

where ρ is the euclidean distance and $\rho(x, A) = \inf_{y \in A} \rho(x, y)$. Then

$$v(\Gamma^r A) \geqslant v(A) + r, \tag{2.7}$$
$$v(\Gamma_r A) \leqslant r - v(A), \tag{2.8}$$
$$v(\Gamma_{-r} A) \leqslant v(A) - r \tag{2.9}$$

with equalities if and only if A is a sphere up to a null set.

(2.7) and (2.8) can be put into another equivalent form. For this define

$$D(A, B) = \sup_{p \in A, q \in B} \rho(p, q), \qquad D^*(A, B) = \inf_{p \in \bar{A}, q \in B} \rho(p, q).$$

Then

$$v(A) - v(B) \geqslant D^*(A, B), \qquad A \supset B, \tag{2.10}$$
$$v(A) + v(B) \leqslant D(A, B) \tag{2.11}$$

with equalities exactly when A, B are concentric spheres up to null sets.

To see this set $r = D^*(A, B)$ resp. $r = D(A, B)$. In the special case $A = B$ (2.11) gives

$$v(A) \leqslant \tfrac{1}{2} D(A, A) \tag{2.12}$$

that is, for given diameter the sphere has maximal volume [4]. We shall see below that all inequalities (2.7)–(2.12) have analogues in Hamming spaces. The exact formulations do not translate because slight modifications are necessary if v is not an integral. Also in order to get exact results for quasi-spheres one has to cope with the fact that boundaries of sets in H^n have non-zero "volume". We use the notion of a volume radius only heuristically. For the exact formulation of the results it is simpler to avoid that notion altogether.

For $x \in H^n$ denote by $\bar{x}$ the element of H^n which satisfies $x + \bar{x} = (1, 1, \ldots, 1) = \mathbf{1}$.

Lemma 2.1. (*Spherical duality of* H^n)

$$\min_{A \subset H^n, |A| = N} \left| \bigcup_{x \in A} S_r(x) \right| = 2^n - \max_{A \subset H^n, |A| = N} \left| \bigcap_{x \in A} S_{n-r}(\bar{x}) \right|.$$

That is, minimizing the union of a given number of spheres with identical radii is equivalent to maximizing the intersection of a given number of spheres with identical radii.

Proof.

$$\min_{A,|A|=N} \left| \bigcup_{x\in A} S_r(x) \right| = \min_{A,|A|=N} \left| \overline{\left(\bigcap_{x\in A} \overline{S_r(x)} \right)} \right|$$

$$= \min_{A,|A|=N} \left| 2^n - \bigcap_{x\in A} S_{n-r}(\bar{x}) \right|$$

$$= 2^n - \max_{A,|A|=N} \left| \bigcap_{x\in A} S_{n-r}(\bar{x}) \right| .$$

Analogously to (2.6) define now for $A \subset H^n$

$$\Gamma' A = \{x : x \in H^n,\ d(x, A) \leqslant r\}, \tag{2.13}$$

$$\Gamma_r A = \{x : x \in H^n,\ d(x, a) \leqslant r \quad \forall a \in A\},$$

$$\Gamma_{-r} A = \{x : x \in H^n,\ d(x, \bar{A}) \geqslant r\}.$$

Clearly,

$$\Gamma' A = \bigcup_{x\in A} S_r(x), \qquad \Gamma_r A = \bigcap_{x\in A} S_r(x), \tag{2.14}$$

$$\Gamma_{-r} A = \{x : d(x, a) \geqslant r \quad \forall a \in \bar{A}\}$$

$$= \{x : d(x, a) \leqslant n - r \quad \forall a \in \mathbf{1} + \bar{A}\}$$

$$= \bigcap_{a\in 1+\bar{A}} S_{n-r}(a).$$

The isoperimetric property and Lemma 2.1 imply

Lemma 2.2. *The quasi-sphere (defined after Theorem 1.1) is a solution to the following extremal problems*

(a) $$\min_{\substack{A\subset H^n \\ |A|=N}} |\Gamma' A| = \min_{A,|A|=N} \left| \bigcup_{x\in A} S_r(x) \right| ,$$

(b) $$\max_{A,|A|=N} |\Gamma_r A| = \max_{A,|A|=N} \left| \bigcap_{x\in A} S_r(x) \right| ,$$

(c) $$\max_{A,|A|=N} |\Gamma_{-r} A| = \max_{A,|A|=N} \left| \bigcap_{a\in 1+\bar{A}} S_{n-r}(a) \right| .$$

We derive now analogously to (2.10) and (2.11),

$$\max_{\substack{A,B,|A|=N \\ D(A,B)\leqslant d}} |B| = \max_{A:|A|=N} \max_{B,D(A,B)\leqslant d} |B| \tag{2.15}$$

$$= \max_{A:|A|=N} |\Gamma_d A| = \max_{A:|A|=N} \left| \bigcap_{x\in A} S_d(x) \right| .$$

By Lemma 2.2 the maximum is assumed if A is a quasi-sphere. The corresponding maximal B is $B = \Gamma_d A$. If $|A| > \sum_{e=1}^{d} \binom{n}{e}$, then $\Gamma_d A = \emptyset$, otherwise $\Gamma_d A$ is also a quasi-sphere. Similarily,

$$\max_{|A|=N} \max_{\substack{D^*(A,B) \geqslant d^* \\ A \supset B}} |B| = \max_{|A|=N} |\Gamma_{-d^*} A| \tag{2.16}$$

and again by Lemma 2.2 the quasi-sphere gives a solution.

We state our main result, (2.15), in set theoretic language as

Theorem 2.3.[2] *Let $1 \leqslant N \leqslant 2^n$, $1 \leqslant d \leqslant n$, and let $\mathfrak{A}$, $\mathfrak{B}$ denote families of subsets of $\{1, \ldots, n\}$. We call $(\mathfrak{A}, \mathfrak{B})$ a d-pair if $|A \,\Delta\, B| \leqslant d$ for all $A \in \mathfrak{A}$, $B \in \mathfrak{B}$. Then*

$$\max_{\substack{d\text{-pairs}(\mathfrak{A},\mathfrak{B}) \\ |\mathfrak{A}|=N}} |\mathfrak{B}|$$

is assumed if $\mathfrak{A}$ is a quasi-sphere and

$$\mathfrak{B} = \Gamma_d \mathfrak{A} = \{B : |B \,\Delta\, A| \leqslant d \quad \forall A \in \mathfrak{A}\}.$$

It can be seen from Lemma 2.1 that this result is equivalent to the isoperimetric property.

In generalizing the Erdös–Ko–Rado Theorem [11] Kleitman proved another two family result.

Theorem 2.4 [21]. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be two families of different subsets of $\{1, \ldots, n\}$, such that $|A| = k$ for all $A \in \mathfrak{A}$, $|B| = l$ for all $B \in \mathfrak{B}$ and*

$$|A \,\Delta\, B| \leqslant k + l - 1 \quad A \in \mathfrak{A},\ B \in \mathfrak{B}.$$

Then either

$$|\mathfrak{A}| \leqslant \binom{n-1}{k-1} \quad \text{or} \quad |\mathfrak{B}| \leqslant \binom{n-1}{l-1}.$$

The Hamming analogue to (2.12) exists already in the literature and has actually been proved twice [17, 20]. In both papers it is shown first that a "pseudo-sphere" is a solution and then the exact boundary is determined with the help of the Erdös–Ko–Rado Theorem [11]. We show below that the first part is an immediate consequence of Theorem 2.3. For the determination of the boundary we need here Theorem 2.4, which is a generalization of the Erdös–Ko–Rado Theorem and which fortunately has also a nice and perspicuous proof [21].

In [20] the problem was formulated as follows: What is the maximal size of a family $\mathfrak{A}$ of subsets of $\{1, \ldots, n\}$ subject to the condition:

$$|A \,\Delta\, B| \leqslant d \quad \text{for all } A, B \in \mathfrak{A}. \tag{2.17}$$

[2] Recently P. Frankl informed us that he has obtained related results. His exact statement is not available to us at the present time.

If one replaces the condition (2.17) by

$$|A \cup B| \leq d \quad \text{for all } A, B \in \mathfrak{A} \tag{2.18}$$

one gets the problem solved in [17].

The equivalence of the two problems is immediate from the observation that one can limit onself in both cases to *order ideals*, that is, families of subsets, which contain with every subset *all* its subsets. (See Lemma 2.6 below.)

Theorem 2.5. *Let $\mathfrak{A}$ be a family of subsets of $\{1, \ldots, n\}$ satisfying for $d \leq n-2$ $|A \Delta B| \leq d$ for all $A, B \in \mathfrak{A}$, then* $\max |\mathfrak{A}|$ *is assumed in case $d = 2r$ by $\mathfrak{A} = \{A : |A| \leq r\}$ and in case $d = 2r+1$ by $\mathfrak{A} = \{A : |A| \leq r\} \cup \{A : |A| = r+1, x \in A$ for a fixed $x, 1 \leq x \leq n\}$.*

Proof. One has to find a family for which $f_d = \max_{\Gamma_d \mathfrak{A} \supset \mathfrak{A}} |\mathfrak{A}|$ is assumed.

Consider more generally

$$g_d = \max_{\substack{(\mathfrak{A}, \mathfrak{B})-d\text{-pair} \\ |\mathfrak{B}| \geq |\mathfrak{A}|}} |\mathfrak{A}| \geq f_d. \tag{2.19}$$

For fixed $\mathfrak{A}$ a best choice for $\mathfrak{B}$ is $\mathfrak{B} = \Gamma_d \mathfrak{A}$. Whatever the cardinality of an optimal $\mathfrak{A}$ may be, by Theorem 2.3 we know that $|\mathfrak{B}| = |\Gamma_d \mathfrak{A}|$ is maximal if $\mathfrak{A}$ is chosen as a quasi-sphere:

$$S_l(\emptyset) \subseteq \mathfrak{A} \subsetneq S_{l+1}(\emptyset). \tag{2.20}$$

Here $S_1(\emptyset)$ denotes a sphere of radius l around the empty set $\emptyset$.

Case 1: $d < 2l$. Then $|\Gamma_d \mathfrak{A}| < |\mathfrak{A}|$, because $\Gamma_d \mathfrak{A}$ cannot contain a set of size $\geq l$, and therefore this does not occur.

Case 2: $d \geq 2l+2$. This contradicts the definition of l and the maximality of $|\mathfrak{A}|$.

Case 3: $d = 2l$. Then $A \in \Gamma_d \mathfrak{A}$ implies $|A| \leq 1$ and the solution is $\Gamma_d \mathfrak{A} = \mathfrak{A} = S_l(\emptyset)$.

Case 4: $d = 2l+1$. Then $S_l(\emptyset) \subset \Gamma_d \mathfrak{A} \subset S_{l+1}(\emptyset)$. By Theorem 2.4 either

$$|\mathfrak{A} - S_l(\emptyset)| \leq \binom{n-1}{l} \quad \text{or} \quad |\Gamma_d \mathfrak{A} - S_l(\emptyset)| \leq \binom{n-1}{l}$$

and since $|\Gamma_d \mathfrak{A}| \geq |\mathfrak{A}|$ certainly $|\mathfrak{A} - S_l(\emptyset)| \leq \binom{n-1}{l}$. By choosing $\mathfrak{A} - S_l(\emptyset)$ in the Erdös–Ko–Rado fashion we get the $\mathfrak{A}$ for which g_d is assumed. In both possible cases $g_d = f_d$ and hence the theorem.

We conclude this section by showing that in both, Theorem 2.3 and Theorem 2.5, the operation "Δ" can be replaced by "$\cup$". The argument seems to have been used for the first time in [11], then in [17] and decisively in [20] and [21].

Lemma 2.6. *Let $1 \leq N$, $M \leq 2^n$, $X = \{1, \ldots, n\}$, $\mathcal{P}(X) =$ power set of X, and $1 \leq d \leq n$.*

(a) *The following two conditions are equivalent*:
There exists $\mathfrak{A}, \mathfrak{B} \subset \mathcal{P}(X)$ *such that* $|\mathfrak{A}| = N$,

$$|\mathfrak{B}| = M \quad \text{and} \quad |A \,\Delta\, B| \leqslant d \quad \text{for all} \quad A \in \mathfrak{A},\ B \in \mathfrak{B}. \tag{2.21}$$

There exists $\mathfrak{A}, \mathfrak{B} \subset \mathcal{P}(X)$ *such that* $|\mathfrak{A}| = N$,

$$|\mathfrak{B}| = M \quad \text{and} \quad |A \cup B| \leqslant d \quad \text{for all} \quad A \in \mathfrak{A},\ B \in \mathfrak{B}. \tag{2.22}$$

(b) *Equivalent conditions are*:
There exists an $\mathfrak{A} \subset \mathcal{P}(X)$ *such that* $|\mathfrak{A}| = N$

$$\text{and} \quad |A \,\Delta\, B| \leqslant d \quad \text{for all} \quad A, B \in \mathfrak{A}. \tag{2.23}$$

There exists an $\mathfrak{A} \subset \mathcal{P}(X)$ *such that* $|\mathfrak{A}| = N$

$$\text{and} \quad |A \cup B| \leqslant d \quad \text{for all} \quad A, B \in \mathfrak{A}. \tag{2.24}$$

Proof. In all four cases one can limit oneself to order ideals. This is obvious for (2.22) and (2.24). In order to see it for the other cases define for $x \in \{1, \ldots, n\}$

$$\mathfrak{A}_x = \{A : A \in \mathfrak{A},\ x \in A,\ A - \{x\} \notin \mathfrak{A}\},$$

$$\mathfrak{B}_x = \{B : B \in \mathfrak{B},\ x \in B,\ A - \{x\} \notin \mathfrak{B}\},$$

$$\mathfrak{A}_x^* = \{A - \{x\} : A \in \mathfrak{A}_x\},$$

$$\mathfrak{B}_x^* = \{B - \{x\} : B \in \mathfrak{B}_x\}$$

and set

$$\mathfrak{A}^* = (\mathfrak{A} - \mathfrak{A}_x) \cup \mathfrak{A}_x^*,$$

$$\mathfrak{B}^* = (\mathfrak{B} - \mathfrak{B}_x) \cup \mathfrak{B}_x^*,$$

then $|\mathfrak{A}^*| = |\mathfrak{A}|$, $|\mathfrak{B}^*| = |\mathfrak{B}|$ and $(\mathfrak{A}^*, \mathfrak{B}^*)$ and $(\mathfrak{A}^*, \mathfrak{A}^*)$ are d-pairs. The claim follows by iteratively applying this to all $x \in \{1, \ldots, n\}$.

If now $\mathfrak{A}$ and $\mathfrak{B}$ are order ideals (especially if $\mathfrak{A} = \mathfrak{B}$), then $|A \,\Delta\, B| \leqslant d$ for all $A \in \mathfrak{A}$, $B \in \mathfrak{B}$ implies that also $|A \cup B| \leqslant d$ for all $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, because $A - B \in \mathfrak{A}$. The converse implication is obvious.

3. Sum type extremal problems

In estimating conditional probabilities for correlated independent processes optimization problems arise which involve functions depending on all pairwise distances of the elements of a set. In order to understand the nature of such problems we consider there a simpler type of such problems without worrying at this time about possible applications. We ask the following questions: What is the structure of a family of subsets of $\{1, \ldots, n\}$ for which

$$\min_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_{A,B\in\mathfrak{A}} |A \Delta B| \tag{1}$$

$$\max_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_{A,B\in\mathfrak{A}} |A \Delta B| \tag{2}$$

$$\min_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_{A,B\in\mathfrak{A}} |A \cap B| \tag{3}$$

$$\max_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_{A,B\in\mathfrak{A}} |A \cap B| \tag{4}$$

is assumed?

Problems involving the union can be transformed into a problem involving intersections by complementation. Similar questions can be asked for several families of sets and for families with a size limitation on the subsets. (2) is actually trivial, we have solved (3), we have a conjecture about (4) and no idea about (1) except that there is some connection between the two.

In order to find a solution to (2) define $\mathfrak{A}_x = \{A : A \in \mathfrak{A},\ x \in A\}$ for $x = 1, 2, \ldots, n$. Since $\Sigma_{A,B\in\mathfrak{A}} |A \Delta B| = 2\Sigma_x |\mathfrak{A}_x| \, |(\mathfrak{A} - \mathfrak{A}_x)|$ and since the function $g(x) = x(|\mathfrak{A}| - x)$ takes its integer valued maximum at $(\frac{1}{2}|\mathfrak{A}|)$, we have to find a family $\mathfrak{A}$ satisfying $|\mathfrak{A}_x| = \frac{1}{2}N$ if N is even and

$$|\mathfrak{A}_x| = [\tfrac{1}{2}N] \quad \text{or} \quad [\tfrac{1}{2}N] + 1 \qquad \text{for } x = 1, \ldots, n \text{ if } N \text{ is odd.}$$

In the first case for instance any family $\mathfrak{A}$ with $A \in \mathfrak{A}$ implies $\bar{A} \in \mathfrak{A}$ is a solution. In the second case choose an $\mathfrak{A}^*$ with $A \in \mathfrak{A}^*$ implies $\bar{A} \in \mathfrak{A}^*$ and $|\mathfrak{A}^*| = N - 1$ and define $\mathfrak{A} = \mathfrak{A}^* \cup \{A\}$, where $A \notin \mathfrak{A}^*$ and is arbitrary otherwise. Of course there are many other solutions.

We state now the solution to problem (3) as

Theorem 3.1. $\min_{\mathfrak{A}:|\mathfrak{A}|=N} \left| \Sigma_{A,B\in\mathfrak{A}} |A \cap B| \right|$ *is assumed for a (pseudo)-sphere characterized by the property that*

$$\big||\{A : A \in \mathfrak{A},\ x \in A\}| - |\{A : A \in \mathfrak{A},\ y \in A\}|\big| \leq 1 \quad \textit{for all } x, y \in \{1, 2, \ldots, n\}.$$

Proof. For $\mathfrak{A} = \{A_1, \ldots, A_N\}$ define the incidence matrix $I = (I_{ij})_{j=1,\ldots,n}^{i=1,\ldots,N}$ by $I_{ij} = 1$ iff $j \in A_i$. $|\mathfrak{A}_x| = |\{A : A \in \mathfrak{A},\ x \in A\}|$ counts the number of 1's in the xth column of I. We have

$$\sum_{A,B\in\mathfrak{A}} |A \cap B| = \sum_x |\mathfrak{A}_x|^2. \tag{3.1}$$

The function $f(x) = x^2$ has the property that for two natural numbers x and y, $x > y$,

$$f(x-1) + f(y+1) \leq f(x) + f(y) \quad \text{with equality iff } x = y + 1. \tag{3.2}$$

Therefore we can decrease $\Sigma_x |\mathfrak{A}_x|^2$ by subtracting from a big column and adding

it to a small column. We show now that this can be done in such a way that the resulting rows are still distinct and hence we get a new family $\mathfrak{A}'$ with $|\mathfrak{A}'| = N$.

Let us suppose that there exists a pair (x, y) with $|\mathfrak{A}_x| \geq |\mathfrak{A}_y| + 2$. Consider now those rows where the xth and the yth column differ. Write the rows which have a 1 in the xth column as

$$a_s\, 10, \qquad s = 1, \ldots, S, \tag{3.3}$$

and those which have a 1 in the yth column as

$$b_t\, 01, \qquad t = 1, \ldots, T. \tag{3.4}$$

Since $S \geq T + 2$ there exists an element a_1, say, such that $a_1 \notin \{b_t \mid t = 1, \ldots, T\}$. Replace now $a_1\, 10$ by $a_1\, 01$. We can iterate the procedure until for all x

$$|\mathfrak{A}_x| = l \quad \text{or} \quad l + 1. \tag{3.5}$$

Given $\mathfrak{A}$ and thus $\Sigma\, |\mathfrak{A}_x|$ it is clear how l is defined. It is the largest integer such that

$$ln \leq \sum_x |\mathfrak{A}_x|. \tag{3.6}$$

There are k, $0 \leq k < n$, terms with values $l + 1$. Since

$$\sum_x |\mathfrak{A}_x| = \sum_{A \in \mathfrak{A}} |A|, \tag{3.7}$$

a decrease in $\Sigma_{A \in \mathfrak{A}} |A|$ can only have the effect that the number of terms with value $l + 1$ decreases, if possible, to 0, then the number of terms with value l decreases and so on. This implies that also $\Sigma_x |\mathfrak{A}_x|^2$ decreases and the procedure stops when $\Sigma_{A \in \mathfrak{A}} |A|$ is minimal, that is for the pseudo-sphere, which is balanced:

$$|\mathfrak{A}_x| = t \quad \text{or} \quad t + 1.$$

It is conceivable that the answer to (4) is a sphere around X, but since the convexity of $f(x) = x^2$ does not help in case of maximisation a completely new and likely harder argument is needed.

We conclude this section with bounds for (4), which are in a certain sense asymptotically sharp. Since

$$f_n(N) = \max_{\mathfrak{A} : |\mathfrak{A}| = N} \sum_x |\mathfrak{A}_x|^2 = |\mathfrak{A}|^2 \max_{\mathfrak{A} : |\mathfrak{A}| = N} \sum_{x=1}^{n} \left(\frac{|\mathfrak{A}_x|}{|\mathfrak{A}|} \right)^2$$

can be interpreted as follows: define a probability distribution on H^n by putting

$$P^n(x^n) = \begin{cases} \dfrac{1}{|\mathfrak{A}|} & \text{for } x^n \in \mathfrak{A}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.8}$$

Then $(|\mathfrak{A}_x|/|\mathfrak{A}|,\ 1 - |\mathfrak{A}_x|/|\mathfrak{A}|)$ is the 1-dimensional marginal distribution on the xth component. $P_x = |\mathfrak{A}_x|/|\mathfrak{A}|$ is the probability for 1 and $1 - p_x$ the probability for 0. By allowing in the "max" general probability distributions on H^n with given entropy — the substitute for cardinality — we get a function

$$g_n(N) = N^2 \max_{P^n : H(P^n)=\log N} \sum_{x=1}^{n} P_x^2 \geq f_n(N). \tag{3.9}$$

It suffices to consider the function

$$G_n(c) = \frac{1}{n} \max_{(1/n)H(P^n)=c} \sum_{i=1}^{n} p_i^2. \tag{3.10}$$

Set $G(c) = G_1(c)$.

Lemma 3.2. *$G(c)$ is monotonically decreasing and concave ($\cap$), $G_n(c) = G(c)$ for $n = 1, 2, \ldots$.*

Proof. $G(c) = p^2$, if $h(p) - p = \log p - (1-p)\log(1-p) = c$ and $p \geq \frac{1}{2}$. Write $G(c) = (h^{-1}(c))^2$.

$$\frac{dG}{dc} = 2h^{-1}(c)\frac{dh^{-1}(c)}{dc} = 2h^{-1}(c)\left(\frac{dc}{dp}\right)^{-1}$$

$$\frac{dc}{dp} = \log\frac{1-p}{p}.$$

Hence

$$\frac{dG}{dc} = 2h^{-1}(c)\left[\log\frac{1-h^{-1}(c)}{h^{-1}(c)}\right]^{-1}, \tag{3.11}$$

$$\begin{aligned}\frac{d^2G}{dc^2} &= 2\left[\log\frac{1-h^{-1}(c)}{h^{-1}(c)}\right]^{-2} + \\ &\quad + 2h^{-1}(c)\left\{-\left[\log\left(\frac{1-h^{-1}(c)}{h^{-1}(c)}\right)\right]^{-2}\left(\frac{h^{-1}(c)}{1-h^{-1}(c)}\right)\right. \\ &\quad \left.\times\left(-\frac{1}{(h^{-1}(c))^2}\left[\log\frac{1-h^{-1}(c)}{h^{-1}(c)}\right]^{-1}\right)\right\} \\ &= 2\left[\log\frac{1-h^{-1}(c)}{h^{-1}(c)}\right]^{-2}\left\{1+\frac{1}{1-h^{-1}(c)}\left[\log\left(\frac{1-h^{-1}(c)}{h^{-1}(c)}\right)\right]^{-1}\right\}.\end{aligned}$$

The first factor is positive. To see that the other bracket is negative it suffices to show that

$$1 \geq -(1-p)\log\frac{1-p}{p}. \tag{3.12}$$

But this is true because

$$-p\log p - (1-p)\log(1-p) \leq 1 \quad \text{and} \quad -\log p > 0.$$

The inequality $G_n(c) \geq G(c)$ is obtained by considering for p^n product distributions.

Since always $H(p^n) \leqslant \sum_{i=1}^n H(p_i)$ we have

$$G_n(c) \leqslant \frac{1}{n} \max_{1/n \sum_{i=1}^n H(p^i) \geqslant c} \sum_{i=1}^n p_i^2 = \max_{1/n \sum_{i=1}^n H(p_i) \geqslant c} \frac{1}{n} \sum_{i=1}^n G(H(p_i))$$

$$\leqslant \max_{1/n \sum_{i=1}^n H(p_i) \geqslant c} G\left(\frac{1}{n} \sum_{i=1}^n H(p_i)\right) \quad \text{(by concavity)}$$

$$= G(c), \text{ since } G \text{ is decreasing.}$$

This completes the proof of Lemma 3.2. We therefore have also proved

$$g_n(N) = N^2 n \left[h^{-1}\left(\frac{\log N}{n}\right)\right]^2. \tag{3.13}$$

For any family $\mathfrak{A}$ containing with every set its complement

$$\sum_{A, B \in \mathfrak{A}} |A \cap B| = \frac{|\mathfrak{A}|^2}{4} \cdot n \leqslant f_n(N).$$

Corollary. $[h^{-1}(\log N/n)]^2 N^2 n \geqslant f_n(N) \geqslant \frac{1}{4} N^2 n$. *For* $N = N(n) \geqslant c \cdot 2^n$, $c > 0$, *this implies*

$$\lim_{n \to \infty} \frac{f_n(N(n))}{N^2 \cdot n} = \frac{1}{4}.$$

4. Optimization for order ideals under a weight assignment

Many extremal problems for families of subsets are such that the answer can be found in the class of families having the order ideal property. The problems dealt with in earlier sections are all of that nature and it makes sense to study order ideals in their own right (see for this also [12]). In [22] the following question is answered: Given a weight-function w_i on the subsets of $\{1, \ldots, n\}$ which depends only on their sizes, what is

$$\max \sum_{i=0}^n K_i(\mathfrak{A}) w_i,$$

if $K_i(\mathfrak{A})$ counts the number of i-tuples in a family of subsets $\mathfrak{A}$ and if the max runs over the families $\mathfrak{A}$ with $A \not\subset B$ for all $A, B \in \mathfrak{A}$?

Another problem arises if the families $\mathfrak{A}$ have the additional property: $A \cap B \neq \emptyset$ for all $A, B \in \mathfrak{A}$. It was solved in [13].

In this section we prove results of this type, when $\mathfrak{A}$ is an order ideal and $|\mathfrak{A}|$ is fixed. Replacing the "max" above by a "min" leads to a trivial problem in the two cases mentioned, however, in our case it is also interesting. Actually, we formulate the problems only for "min", because the case of "max" follows from it by taking the weights $-w_i$. The methods of proof heavily rely on the Kruskal Theorem [23], which was independently obtained in [19]. Meanwhile several elegant proofs exist [6, 8, 9]. It is also at the root of the isoperimetric property and states the following:

Theorem 4.1. *If $\mathfrak{A}$ is a family of k-element subsets of an n-set and $\delta(\mathfrak{A})$ denotes the family of all $(k-1)$-element subsets, which are subsets of a set in $\mathfrak{A}$, then*

$$|\delta(\mathfrak{A})| \geq F(k, m), \tag{4.1}$$

where $m = |\mathfrak{A}|$ and

$$F(k, m) = \binom{a_k}{k-1} + \binom{a_{k-1}}{k-2} + \cdots + \binom{a_t}{t-1},$$

if the k-canonical representation of m is

$$m = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \cdots + \binom{a_t}{t},$$

where $a_k > a_{k-1} > \cdots > a_t \geq t \geq 1$.

Equality holds in (4.1), *when the k-subsets are chosen in lexicographic order. Also, $F(k, m)$ is monotonically increasing in m.*

Corollary. *If $\mathfrak{A}$ is an order ideal on an n-element set and K_i is the number of i-element members in $\mathfrak{A}$, $0 \leq i \leq n$, then*

$$K_{i-1} \geq F(i, K_i), \qquad i = 1, 2, \ldots, n, \tag{4.2}$$

and conversely, if (4.2) *holds for numbers K_i $(i = 0, \ldots, n)$ then there exists an order ideal with K_i members on the ith level.*

Proof. The first implication follows immediately from the order ideal property and Theorem 4.1. To get the second implication simply choose $\mathfrak{A}$ such that for every i $(i = 0, 1, \ldots, n)$ the K_i i-tuples are the first in lexicographic order.

Let now w_i $(i = 0, 1, \ldots, n)$ denote the weight assigned to each i-tuple, we are interested in minimizing or maximizing $\sum_{i=1}^{n} w_i K_i$ for order ideals with level number K_i. We shall impose monotonicity restrictions on the sequence $(w_0, w_1, \ldots, w_n)$.

In order to state our Theorem 4.2 below we need the definition of a quasi-cylinder. If $|\mathfrak{A}| = 2^s$ for some $s, 0 \leq s \leq n$, then all the subsets of an s-element set form a cylinder. We can always assume that the s-element set equals $\{1, 2, \ldots, s\}$. A quasi-cylinder is a generalization of this concept for other cardinalities.

A *quasi-cylinder* of cardinality N consists of the first N subsets of $\{1, \ldots, n\}$ in lexicographic order. Observe that if $N = 2^{b_1} + \cdots + 2^{b_r} (b_1 > \cdots > b_r \geq 0)$ then the quasi-cylinder consists of all the subsets of a b_1-element set B_1, of all the sets of the form $A \cup \{a_1\}$, where $a_1 \notin B_1$ and A is an arbitrary subset of a b_2-element set $B_2 \subset B_1$, of all the sets of the form $A \cup \{a_1\} \cup \{a_2\}$, where $a_2 \in B_1$, $a_2 \notin B_2$ and A is an arbitrary subset of a b_3-element set $B_3 \subset B_2$, and so on. The number of i-element subsets in this quasi-cylinder is

$$\binom{b_1}{i} + \binom{b_2}{i-1} + \binom{b_3}{i-2} + \cdots + \binom{b_r}{i-r+1}.$$

We call $b_1, \ldots, b_r$ the parameters of the quasi-cylinder.

Theorem 4.2. *Denote by* $K_i = K_i(\mathfrak{A})$ $(i = 0, 1, \ldots, n)$ *the number of* i*-element members of an order ideal* $\mathfrak{A}$ *and let* $1 \leqslant N \leqslant 2^n$. *Then*

(a) *for* $w_0 \leqslant w_1 \leqslant w_2 \leqslant \cdots \leqslant w_n$, $1 \leqslant N \leqslant 2^n$ $\min_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_i K_i w_i$ *is assumed if* $\mathfrak{A}$ *is a quasi-sphere.*

(b) *for* $w_0 \geqslant w_1 \geqslant \cdots \geqslant w_n$, $1 \leqslant N \leqslant 2^n$, $\min_{\mathfrak{A}:|\mathfrak{A}|=N} \sum K_i w_i$ *is assumed if* $\mathfrak{A}$ *is a quasi-cylinder.*

Proof. (a) is trivial and stated only for comparison. The proof of (b) is given in two parts, which we state as Lemma 4.3 and Lemma 4.4. In general there are several families for which

$$\min_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_{i=1}^{n} K_i w_i \quad \text{is assumed.}$$

We get uniqueness by allowing only optimal families for which on every level the elements are the first in lexicographic order and in addition

$$\sum_{i=1}^{n} i K_i \quad \text{is maximized.} \tag{4.3}$$

This helps in the proofs. We denote the unique optimal order ideal by $O(N, n, \boldsymbol{w})$, where $\boldsymbol{w} = (w_0, \ldots, w_n)$.

Lemma 4.3. *For the optimal order ideal* $O(N, n, \boldsymbol{w})$ *we have*

(a) $F(i, K_i) \leqslant K_{i-1} \leqslant F(i, K_i + 1)$ $\quad (i = 1, \ldots, n)$

and

(b) *either we have strict inequality for every* i *on the right hand side of* (a) *or there is an* s *with*

$$K_{s-1} = F(s, K_s + 1),\ K_{i-1} < F(i, K_i + 1) \qquad (i = s+1, \ldots, n)$$

and then $F(i, K_i) = K_{i-1}$ $\quad (i = 1, 2, \ldots, s-1)$.

Lemma 4.4. *If for* $\boldsymbol{K} = (K_0, K_1, \ldots, K_n)$ (a) *and* (b) *in Lemma* 4.3 *hold, then one can find integers* $b_1 > b_2 > \cdots > b_r \geqslant 0$ *such that*

$$\text{(c)}\quad K_i = \binom{b_1}{i} + \binom{b_2}{i-1} + \cdots + \binom{b_{j(i)}}{i - j(i) + 1} \qquad (i = 0, 1, \ldots, n),$$

where $j(i) \geqslant r$, $b_j(i) \geqslant i - j(i) + 1$, *but* $b_j(i) + 1 < i - j(i)$. *Moreover* $b_i = s - i$ *if* $j(s) + 1 \leqslant i \leqslant s$.

The number in (c) is exactly the number of i-element members of a quasi-cylinder with parameters $b_1, \ldots, b_r$. Thus, if we choose the first K_i i-tuples in lexicographic order for all i, then we obtain a quasi-cylinder. Therefore Lemmas 4.3 and 4.4 yield (b) of Theorem 4.2.

Proof of Lemma 4.3. (a) Denote by $\boldsymbol{K} = (K_0, K_1, \ldots, K_n)$ the vector of level numbers for $O(n, N, w)$. For p, q; $1 \leq p < q \leq n$; define the transformation $T_{p,q} : \boldsymbol{K} \to \boldsymbol{K}'$ by

$$K'_p = K_p - 1, \quad K'_q = K_q + 1$$

$$K'_i = K_i \quad \text{for } i \neq p, q. \tag{4.4}$$

Clearly, $\sum K'_i w_i \leq \sum K_i w_i$ and $\sum i K'_i > \sum i K_i$. Therefore $\boldsymbol{K}$ cannot be optimal, if $\boldsymbol{K}'$ satisfies (4.2). If $\boldsymbol{K}$ satisfies (4.2) and $\boldsymbol{K}'$ not, then this could have only two reasons because $F(i, K_i)$ is monotonically increasing in K_i:

$$K'_p = K_p - 1 < F(p+1, K'_{p+1}) \quad \text{or}$$

$$K'_{q-1} < F(q, K'_q) = F(q, K_q + 1). \tag{4.5}$$

For $q = p + 1$ those two inequalities are the same and therefore

$$K_p \leq F(p+1, K_{p+1} + 1) \quad \text{for} \quad p = 1, 2, \ldots, n. \tag{4.6}$$

This proves (a).

(b) Suppose that there exists a q with

$$K_{q-1} = F(q, K_q + 1) \tag{4.7}$$

(if there are more choose the largest) and a $p, p + 1 < q$, with

$$F(p+1, K_{p+1}) < K_p. \tag{4.8}$$

Apply the transformation $T_{p,q}$.

Since $K_{p-1} \geq F(p+1, K_{p+1}) = F(p+1, K'_{p+1})$ (4.5) could hold only if

$$K'_{q-1} = K_{q-1} < F(q, K_q + 1), \quad \text{but this contradicts (4.7).}$$

Proof of Lemma 4.4. We proceed by induction in i. Suppose that $K_n = \cdots = K_{m+1} = 0$ and $K_m > 0$. Let us first assume that we are in the case

$$K_{i-1} < F(i, K_i + 1) \qquad (i = 1, 2, \ldots, n). \tag{4.9}$$

We prove the statement first for $i = m$ and then, decreasing always by 1, for all $i \geq 0$. For $i = m + 1$, (4.9) gives $0 \leq K_m < F(m+1, 1) = m + 1$. That means K_m can be written in the m-canonical form

$$K_m = \binom{m}{m} + \binom{m-1}{m-1} + \cdots + \binom{m - K_m + 1}{m - K_m + 1}, \quad \text{where } m - K_m + 1 \geq 1.$$

Observe that $b_1 = w$, $b_2 = w - 1, \ldots, b_{K_m} = m - K_m + 1$.

Suppose now that the statement holds for i and let us prove it for $i - 1$. Thus, $b_1, \ldots, b_j(i)$ are already defined and they satisfy (c). We make use of the fact that

$$F(i, K+1) - F(i, K) = t - 1 \quad \text{if } K = \binom{a_i}{i} + \cdots + \binom{a_t}{t}. \tag{4.10}$$

This and (4.9) imply

$$0 \leqslant K_{i-1} - F(i, K_i) = z = i - j(i) - 1. \tag{4.11}$$

Here

$$F(i, K_i) = \binom{b_1}{i-1} + \cdots + \binom{b_{j(i)}}{i-j(i)} \tag{4.12}$$

and therefore the $(i-1)$-canonical form of K_{i-1} is

$$\binom{b_1}{i-1} + \cdots + \binom{b_{j(i)}}{i-j(i)} + \binom{i-j(i)-1}{i-j(i)-1} + \binom{i-j(2)-2}{i-j(2)-2} + \cdots + \binom{i-j(i)-z}{i-j(i)-z}.$$

If $z = 0$, the canonical form of K_{i-1} is given by (4.12), where again $i - j(i) - z \geqslant 1$, by (4.11), and the lemma is proved in this case even in a slightly stronger form: $b_r > 0$. This is important for the remaining case, which we now consider:

$$K_{s-1} = F(s, K_s + 1),\ K_{i-1} < F(i, K_i + 1), \qquad i = s+1, \ldots, n,$$
$$F(i, K_i) = K_{i-1}, \qquad i = 1, 2, \ldots, s-1. \tag{4.13}$$

We can prove in exactly the same way as earlier that (c) holds for all $i \geqslant s$ with suitable b's. Moreover, by the above remark $s \geqslant j(s)$, that is, (c) is still an s-canonical form of K_s. For the next step we have from (4.10) and (4.13)

$$K_{s-1} = F(s, K_s) + s - j(s) \tag{4.14}$$

and therefore

$$K_{s-1} = \binom{b_1}{s-1} + \cdots + \binom{b_{j(s)}}{s-j(s)} + \binom{s-j(s)-1}{s-j(s)-1} + \cdots + \binom{1}{1} + \binom{0}{0}. \tag{4.15}$$

This is not an $(s-1)$-canonical form because of the term $\binom{0}{0}$. However, using (4.10) we obtain

$$K_{s-2} = F(s-1, K_{s-1}) = F(s-1, K_{s-1} - 1) + (1-1) = F(s-1, K_{s-1} - 1)$$
$$= \binom{b_1}{s-2} + \cdots + \binom{b_{j(s)}}{s-j(s)-1} + \binom{s-j(s)-1}{s-j(s)-2} + \cdots + \binom{1}{0}. \tag{4.16}$$

By continuing in the same way:

$$K_i = \binom{b_1}{i} + \cdots + \binom{b_{j(s)}}{i-j(s)+1} + \binom{s-j(s)-1}{i-j(s)} + \cdots + \binom{s-i-1}{0}$$
$$(j(s) \leqslant i \leqslant s-1)$$

and

$$K_i = \binom{b_1}{i} + \cdots + \binom{b_{i+1}}{0} \qquad (0 \leqslant i < j(s)).$$

We proved the existence of b's:

$$b_i = s - i \quad \text{if } j(s) + 1 \leqslant i \leqslant s.$$

Theorem 4.5. *Denote by* $K_i = K_i(\mathfrak{A})$ $(i = 0, 1, \ldots, n)$ *the number of* i*-element members of an order ideal. Then*

$$\min_{\mathfrak{A}:|\mathfrak{A}|=N} \sum_{i=1}^{n} K_i w_i, \qquad 1 \leq N \leq 2^n,$$

is assumed

(a) *in case* $w_0 \leq w_1 \leq \cdots \leq w_M \geq w_{M+1} \geq \cdots \geq w_n$ *by a union of a quasi-cylinder and a quasi-sphere,*

(b) *in case* $w_0 \geq w_1 \geq \cdots \geq w_M \leq w_{M+1} \leq \cdots \leq w_n$ *by an intersection of a quasi-cylinder and a quasi-sphere.*

Proof. (a) As in the proof of Theorem 4.2 choose the first K_i i-tuples in lexicographic order, and if there are more systems minimizing $\sum K_i w_i$ then choose one for which in addition

$$\sum K_i \,|M - i| \quad \text{is maximal.}$$

We can repeat the steps of the proof of Lemma 4.3 for values $i \geq M + 1$. That is (a) and (b) of Lemma 4.3 hold for $i \geq M + 1$.

Furthermore we claim that for values of $i \leq M$ an optimal $\boldsymbol{K}$ satisfies:

$$\begin{aligned} K_i &= F(i+1, K_{i+1}) && (\rho + 1 < i \leq M) \\ K_i &= \binom{n}{i} && (0 \leq i \leq \rho) \\ K_{\rho+1} &\geq F(\rho + 2, K_{\rho+2}) && (4.17) \end{aligned}$$

for some $\rho \leq 0$.

To see this suppose that for (p, q); $p < q \leq M$;

$$K_q > F(q+1, K_{q+1}) \quad \text{and} \quad \binom{n}{p} > K_p \tag{4.18}$$

and assume that q is maximal and p is minimal with this property.

Then the transformation $T_{q,p}$:

$$K_i' = K_i (i \neq p, q), \quad K_q' = K_q - 1, \quad K_p' = K_p + 1$$

does not increase $\sum K_i w_i$ and increases $\sum K_i \,|M - i|$. Also, $K_{i-1}' \geq F(i, K_i')$ for $i = 1, \ldots, M$, where the only critical case $K_{p-1} \geq F(p, K_p + 1)$ follows from $K_{p-1} = \binom{n}{p-1}$. Thus (4.18) would imply that $\boldsymbol{K}$ could not be optimal. The negation of (4.18) is (4.17). Let us now choose the first K_i^c i-tuples in lexicographic order for all i's $(0 \leq i \leq n)$, where

$$\begin{aligned} K_i^c &= K_i && (\rho + 1 < i \leq n) \\ K_i^c &= F(i+1, K_{i+1}^c) && (0 \leq i \leq \rho + 1). \end{aligned} \tag{4.19}$$

The system of K_i^c's satisfy (b) of Lemma 4.3. By Lemma 4.4 this implies that the corresponding family $\mathfrak{A}_c$ is a quasi-cylinder. On the other hand let us choose the first K_i^s i-tuples in lexicographic order, where

$$K_i^s = 0 \qquad (\rho + 1 < i \leqslant n)$$

$$K_i^s = K_i \qquad (0 \leqslant i \leqslant \rho + 1). \tag{4.20}$$

This defines (a) quasi-sphere $\mathfrak{A}_s$ and clearly $\mathfrak{A} = \mathfrak{A}_s \cup \mathfrak{A}_c$.

(b) For $0 \leqslant i \leqslant M$ (a) and (b) of Lemma 4.3 hold. Set $K_i^c = K_i (0 \leqslant i \leqslant M)$. Let us define the numbers $K_i^c (M \leqslant i \leqslant n)$ inductively: $K_M^c = K_M$, K_{i+1}^c is the largest number satisfying

$$K_i^c \geqslant F(i+1, K_{i+1}^c). \tag{4.21}$$

By definition these numbers satisfy $K_i^c < F(i+1, K_{i+1}^c + 1)$ and therefore $\boldsymbol{K}^c$ satisfies (a) and (b) of Lemma 4.3. By Lemma 4.4 the corresponding family $\mathfrak{A}_c$ is a cylinder. We prove now that for an optimal $\boldsymbol{K}$:

$$K_i = K_i^c \qquad (M \leqslant i \leqslant \rho)$$

$$K_i = 0 \qquad (\rho + 1 < i \leqslant n)$$

for some $\rho \geqslant M$.

This will complete the proof, because then our optimal family is the intersection of $\mathfrak{A}_c$ and the quasi-sphere $\mathfrak{A}_s$ given by

$$K_i^s = \binom{n}{i} \qquad (0 \leqslant i \leqslant \rho)$$

$$K_{\rho+1}^s = K_{\rho+1}.$$

Suppose that (4.22) does not hold, thus for some (p, q), $M \leqslant p < q \leqslant n$,

$$0 < K_p \neq K_p^c, \qquad K_q > 0$$

and also

$$K_p < K_p^c \quad \text{(by (4.21))}. \tag{4.23}$$

Let p be minimal and let q be maximal with property (4.23), then

$$K_i = K_i^c (M \leqslant i < p) \quad \text{and} \quad K_{q+1} = 0.$$

The transformation $T_{q,p}$:

$$K_q' = K_{q-1}, \qquad K_p' = K_{p+1}$$

would improve our family and using (4.21) one readily verifies that (4.2) still holds. The proof is complete.

Remark 1. In Theorem 4.5 we did not determine "the" best family, we proved only that it must have a certain pattern.

Remark 2. A theorem of Lindsey says (see [7, 15, 24, 3])

$$\max_{|\mathfrak{A}|=N} |\{(A, B): A, B \in \mathfrak{A},\ d(A, B) = 1\}| \tag{4.24}$$

is assumed for a quasi-cylinder. This a consequence of our Theorem 4.2. To see this we have to prove first that there is an optimal $\mathfrak{A}$ which is an order ideal. This can be seen by applying the transformation, which omits a fixed element a from all the subsets $A \in \mathfrak{A}$ such that $A - \{a\} \notin \mathfrak{A}$. Next, for an order ideal $\mathfrak{A}$

$$|\{(A, B): A, B \in \mathfrak{A},\ A \subset B,\ |A| = i - 1,\ |B| = i\}| = K_i(\mathfrak{A})i$$

and the expression in (4.24) equals

$$\max_{\substack{\mathfrak{A} \text{ order ideal} \\ |\mathfrak{A}|=N}} \sum_{i=1}^{n} K_i(\mathfrak{A})i = -\min \sum_{i=1}^{n} K_i(\mathfrak{A})(-i).$$

Theorem 4.2(b) yields that a quasi-cylinder is an optimal $\mathfrak{A}$.

5. An application to random graphs

Let G be a non-directed graph. The edges of G are deleted independently with probability p, and it is asked whether the new graph will or will not possess a certain property. For instance, if G is k-times connected, what is the probability that the new graph G′ is not k-times connected. In this case the graphs G′, which are not k-times connected, form an order ideal as a family of subsets of the edges of G. We shall not consider a specific graphic property, we give estimates in general, where only one thing is assumed for the property in question: if a G′ does not have the given property and G″ is a subgraph of G′ (subsets of the edges), then G″ does not have the property either.

In other words we have the sequence $11\cdots1$ of length n (the number of edges) and an order ideal $\mathfrak{A}$ is given. Our aim is to give estimates on the probability $P(\mathfrak{A})$ of the event, that $11\cdots1$ goes to $\mathfrak{A}$, if the 1's can change independently to 0 with probability p.

If K_i denotes the number of i-element subsets in $\mathfrak{A}$, then

$$P(\mathfrak{A}) = \sum_{i=0}^{n} K_i p^{n-i}(1-p)^i. \tag{5.1}$$

Here $p^{n-i}(1-p)^i$ is a monotonically increasing or decreasing function of i depending on whether $p < \frac{1}{2}$ or $p > \frac{1}{2}$. By Theorem 4.2, Section 4, fixing the size of $|\mathfrak{A}|$ we obtain lower and upper estimates on $P(\mathfrak{A})$.

If we are interested in how $P(\mathfrak{A})$ changes with p, estimates on the derivative of $P(\mathfrak{A})$ are useful.

$$\frac{dP(\mathfrak{A})}{dp} = \sum_{i=0}^{n} K_i[(n-i)p^{n-i-1}(1-p)^i - p^{n-i}i(1-p)^{i-1}]$$

Here

$$(n-i)p^{n-i-1}(1-p)^i - p^{n-i}i(1-p)^{i-1}$$
$$\leqslant (n-i-1)p^{n-i-2}(1-p)^{i+1} - (i+1)p^{n-i-1}(1-p)^i$$

holds if and only if

$$i \geqslant \frac{2np^2 - 3np + p + n - 1}{1-2p} .$$

That is, these coefficients are decreasing until a certain point and from this point on they are increasing. We can use Theorem 4.5 to get lower and upper estimates.

References

[1] R. Ahlswede and J. Körner, Source coding with side information and a converse for degraded broadcast channels, IEEE Trans. Information Theory 21 (1975) 592–637.

[2] R. Ahlswede, P. Gáces and J. Körner, Bounds on conditional probabilities and applications in multiuser communication, Z. Wahrscheinlichkeitstheorie und verw. Geb. 34 (1976) 157–177.

[3] A.J. Bernstein, Maximally connected arrays on the n-cube, SIAM J. Appl. Math. 15 (1967) 1485–1489.

[4] L. Bieberbach, Über eine Extremaleigenschaft des Kreises. Jber. dtsch. Math. Ver. 24 (1915) 247–250.

[5] H. Brunn, Über Ovale und Eiflächen, Inaugural Diss., München (1887).

[6] G.F. Clements and B. Lindström, A generalization of a combinatorial theorem of Macaulay, J. Combinatorial Theory 7 (1969) 230–238.

[7] G.F. Clements, Sets of lattice points which contain a maximal number of edges, Proc. Am. Math. Soc. 27 (1971) 13–15.

[8] D.E. Daykin, A simple proof of the Kruskal–Katona theorem, J. Combinatorial Theory 17 (1974) 252–253.

[9] J. Eckhoff and G. Wegner, Über einen Satz von Kruskal, Periodica, to appear.

[10] P. Erdös and D.J. Kleitman, Extremal problems among subsets of a set, Proc. Second Chapel Hill Conf. on Combinatorial Mathematics and its Applicatcns, Chapel Hill, N.C. (1970) 146–170.

[11] P. Erdös, Ko. Chao and R. Rado, Intersection theorems for systems of finite sets, Quart. J. Math. Oxford Ser. 12 (1961) 48.

[12] C. Greene and D.J. Kleitman, Proof Techniques in the Theory of Finite Sets, MIT Lecture Notes.

[13] C. Greene, G.O.H. Katona and D.J. Kleitman, Extensions of the Erdös–Ko–Rado Theorem, Submitted to the Proceedings of the Second Prague Conference on Graph Theory.

[14] R.W. Hamming, Error detecting and error correcting codes, Bell System Tech. J. 29 (1950) 147–160.

[15] L.H. Harper, Optimal assignments of numbers to vertices, J. Soc. Industr. Appl. Math. 12 (1964) 131–135.

[16] G.O.H. Katona, The Hamming sphere has minimal boundary, To appear in Studia Sci. Math. Hung.

[17] G.O.H. Katona, Intersection theorems for systems of finite sets, Acta Math. Acad. Sci. Hungar. 15 (1964) 329–337.

[18] G.O.H. Katona, Extremal problems for hypergraphs, Proceedings of the Advanced Study Institute on Combinatorics held at Nijenrode Castle Breukelen, The Netherlands (July 8–20, 1974).

[19] G.O.H. Katona, A theorem of finite sets, Theory of Graphs, Proc. Colloquium held at Tihany 1966 (Akadémiai Kiadó, Budapest, 1968) 187–207.

[20] D. Kleitman, On a combinatorial conjecture of Erdös, J. Combinatorial Theory 1 (1966) 209–214.

[21] D.J. Kleitman, On a conjecture of Milner on K-graphs with non-disjoint edges, J. Combinatorial Theory 5 (1968) 153–156.
[22] D. Kleitman, M. Edelberg and D. Lubell, Maximal sized antichains in partial orders, Discrete Math. 1 (1971) 47–53.
[23] J.B. Kruskal, The number of simplicies in a complex, Mathematical Optimization Techniques (Univ. California Press, Berkeley and Los Angeles, 1963) 251–278.
[24] J.H. Lindsey, Assignment of numbers to vertices, Am. Math. Monthly 71 (1964) 508–516.
[25] F.S. Macaulay, Some properties of enumeration in the theory of modular systems, Proc. London Math. Soc. 26 (1927) 531–555.
[26] G.A. Margulis, Probabilistic properties of graphs with large connectivity, Problemy Peredači Informacii 10 (1974) 101–108.
[27] H. Minkowski, Gesammelte Abhandlungen 2 (Teubner, Leipzig, 1911).
[28] E. Schmidt, Der Brunn-Minkowskische Satz und sein Spiegeltheorem sowie die isoperimetrische Eigenschaft der Kugel in der euklidischen und hyperbolischen Geometrie, Math. Ann. 120 (1948) 307–422.
[29] E. Schmidt, Die Brunn-Minkowskische Ungleichung und ihr Spiegelbild sowie die isoperimetrische Eigenschaft der Kugel in der euklidischen und nicht-euklidischen Geometrie I, Math. Nachr. 1 (1948) 81–157; II. Math. Nachr. 2 (1949) 171–244.
[30] H.A. Schwarz, Beweis des Satzes, daß die Kugel eine kleinere Oberfläche besitzt als jeder andere Körper gleichen Volumens, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. (1884) 1–13.