Association Schemes and Coding Theory

Philippe Delsarte and Vladimir I. Levenshtein

Abstract— **This paper contains a survey of association scheme theory (with its algebraic and analytical aspects) and of its applications to coding theory (in a wide sense). It is mainly concerned with a class of subjects that involve the central notion of the distance distribution of a code. Special emphasis is put on the linear programming method, inspired by the MacWilliams transform. This produces upper bounds for the size of a code with a given minimum distance, and lower bounds for the size of a design with a given strength.**

The most specific results are obtained in the case where the underlying association scheme satisfies certain well-defined "polynomial properties"; this leads one into the realm of orthogonal polynomial theory. In particular, some "universal bounds" are derived for codes and designs in polynomial type association schemes.

Throughout the paper, the main concepts, methods and results are illustrated by two examples that are of major significance in classical coding theory, namely the Hamming scheme and the Johnson scheme. Other topics that receive special attention are spherical codes and designs, and additive codes in translation schemes, including $\mathbb{Z}_4$-additive binary codes.

Keywords— **Association schemes, codes and designs, duality, linear programming, orthogonal polynomials, polynomial schemes, translation schemes, universal bounds.**

I. Introduction

Association scheme theory is part of what is now called algebraic combinatorics [10], [54]. It has two main origins. *Association schemes* were introduced in statistical (combinatorial) design theory by Bose and Shimamoto [23], and the appropriate algebraic setting was given by Bose and Mesner [21]. In fact, the subject can be traced back to a paper by Bose and Nair in 1939 [22].

The second origin is group theory and, more precisely, *character theory of finite groups*, developed by Frobenius, Schur and Burnside. For example, as pointed out by Bannai and Ito [10], a paper by Hoheisel in 1939 derives the orthogonality relations for group characters by a method belonging to "association scheme theory" (before the appearance of association schemes in combinatorics) [61]. Another pioneering contribution in this area is a paper by Kawada on character algebras [67] (see [10]). In fact, one may even say that association scheme theory is as old as Frobenius' representation theory of finite groups (see [11]).

In combinatorics, an association scheme is defined in terms of certain *regularity properties*. In the "group case", the association scheme structure arises from certain *symmetry properties*, which directly induce the desired regularity properties. Thus, following Bannai and Ito, we may say that association scheme theory is a "group theory without groups" [10]. Such a distinction between regularity and symmetry can be found in several subjects. An important example, which belongs to association scheme theory, is the distinction between distance-regular graphs and distance-transitive graphs [26].

The association scheme approach was introduced in *coding theory* in 1973 [37] to deal with a collection of topics involving the notion of the "distance distribution" of a code (see [35], [36]). One of the main subjects is the general concept of a τ-design or a code with "dual distance" $\tau+1$ and a universal (lower) bound on the size of τ-designs. (Term "universal" means here that the bound is valid for all τ-designs in all association schemes under consideration.) This allows one to explain the unified nature of different combinatorial objects and bounds. If a covering radius of a code Y in a metric space X characterizes a degree of the approximation of *any* element of X by elements of Y, then the "dual distance" of Y characterizes an approximation degree of X by Y "at the whole". This idea turned out to be very useful for some problems of numerical analysis [45] and cryptography [122] and was extended to any finite and compact infinite metric spaces in [81]. Another important topic is the problem of finding a universal (upper) bound on the size of a code with minimum distance $\geq d$ or, briefly, a d-code (see [82], [88]). Short introductions to "association schemes and coding theory" were given by Sloane [115] and by Goethals [55]. The same subject is treated in detail in a recent paper by Camion [32].

One of the most significant (although elementary) discoveries was the fact that the MacWilliams transform[1] of the distance distribution of *any* code is non-negative as the mean value of nonnegative definite functions (matrices) over the code [35], [37]. This "innocent appearing result" (to quote Welch, McEliece and Rumsey [131]) has far reaching consequences. The method of obtaining bounds for d-codes based on nonnegative definite functions $F(x,y)$ depending on distance $d(x,y)$, i.e., $F(x,y) = f(d(x,y))$, has been applied by Blichfeldt [19], Rankin [97], Sidelnikov [110], [111]. However, for association schemes (and some of their generalizations) there is a description of all such functions. This makes it possible to apply a linear programming method for finding the best universal bound for d-codes (and τ-designs as well) [35], [37]. For a class of association schemes of important interest for coding theory, the corresponding linear programs can be treated as extremum problems for systems of orthogonal polynomials. Thus, any choice of a permissible polynomial gives rise to a universal bound for d-codes. In 1977, McEliece,

P. Delsarte is with the Department of Computing Science and Engineering, Catholic University of Louvain, Louvain-la-Neuve, Belgium.

V.I. Levenshtein is with the Keldysh Institute for Applied Mathematics, Russian Academy of Science, 125047 Moscow, Russia. His research was supported by the Russian Foundation for Basic Research (grant 98-01-00146) and by the Civilian Research and Development Foundation (grant RM1-346).

[1] Recall that the *MacWilliams identities* relate the weight distribution of a linear code (over a finite field) with that of its orthogonal code by a well-defined linear transform (over the reals) [86], [87].

Rodemich, Rumsey and Welch [90] proposed a polynomial which gives an improvement of the best asymptotic bound obtained before in [111]. One year later, another polynomial was proposed [73]; it gives rise to a universal bound for d-codes that improves upon the MRRW bound and is attained for many cases in different spaces although it gives the same asymptotic result. It turned out [77], [112] that this polynomial is an optimal solution of the corresponding extremum problem in the class of polynomials of a restricted degree. This progress in bounding d-codes allowed one to improve bounds on the *Shannon reliability function* [107] for some probabilistic channels (see [65]).

In classical coding theory, dealing with codes in a *Hamming scheme*, the MacWilliams transform involves a family of orthogonal polynomials [121] known as the *Krawtchouk polynomials* [68]. Surprisingly enough, this fact was not uncovered before 1972 [35], although the "polynomial property" of the MacWilliams transform was pointed out by MacWilliams herself in 1963 [86]. The importance of the role played by Krawtchouk polynomials in coding theory is well recognized nowadays [78], [82], [88]. It can be explained by the fact that these polynomials give the *eigenvalues of the distance relation matrices of the Hamming scheme* [37]. This was first proved implicitly by Vere-Jones in the binary case [128]. A thorough investigation of the group-theoretic significance of the Krawtchouk polynomials was given by Dunkl [46].

The familiar "block codes of length n over a q-ary alphabet", which belong to classical coding theory, can be called "codes in the Hamming (association) scheme H_q^n". The general association scheme approach provides us naturally with a considerable extension of the theory in that it applies to *"codes"* and *"designs"* in *any association scheme* [37], [39]. This combinatorial structure consists of a nonempty finite set X endowed with a collection of binary relations $R_0, R_1, \cdots, R_n$ having strong regularity properties. The adjacency matrices of the graphs (X, R_i) generate a commutative and associative algebra (over the complex numbers) both for the matrix product and the pointwise product. This is called the *Bose–Mesner algebra* of the association scheme. It has two distinguished bases: the basis consisting of the *adjacency matrices* D_i, and the basis consisting of the *irreducible idempotent matrices* E_k. By definition, there exist well-defined complex numbers $p_i(k)$ and $q_k(i)$ such that

$$D_i = \sum_{k=0}^{n} p_i(k) E_k, \quad |X| E_k = \sum_{i=0}^{n} q_k(i) D_i.$$

The *p-numbers* $p_i(k)$ and the *q-numbers* $q_k(i)$ play a prominent role in the theory. They satisfy some well-defined *orthogonality relations.* (In the case of the Hamming scheme, we have $p_i(k) = q_i(k) = K_i(k)$, where $K_i(z)$ is the Krawtchouk polynomial of degree i.) It appears that the p-numbers $p_i(k)$ are the eigenvalues of the adjacency matrix D_i.

There is an important *formal duality* in the theory, called the Krein duality, which permutes the roles of the matrix and pointwise products in the Bose–Mesner algebra [8], [42], [95]. This duality is a rich source of research ideas: "trying to make the theory closed under duality".

A *code* Y in an association scheme is a nonempty subset of the point set X (with the inherited relations $R_i|Y$). The *inner distribution*[2] of Y is the $(n+1)$-tuple $(a_i)_{i=0}^n$ where $|Y|a_i$ counts the ordered pairs of code points $y, y' \in Y$ with $(y, y') \in R_i$. In this general context, the "innocent appearing result" alluded to above is the fact that *the Q-transform of the inner distribution is non-negative,* in the sense that $\sum_{i=0}^n a_i q_k(i)$ is a nonnegative real number (for $k = 0, 1, \cdots, n$). This is the basis of the *linear programming method* to find upper bounds for D-codes and lower bounds for D-designs in an association scheme [37]. "Duality" between D-codes and D-designs manifests itself in the fact that any linear programming bound for D-codes gives a linear programming bound for D-designs and conversely [80]. Explicit universal bounds for codes and designs in some classes of association schemes have been obtained by use of this approach [37], [73], [76], [77], [81].

Certain parts of the theory can be developed further, when appropriate restrictive assumptions are imposed on the p- or q-numbers. An association scheme is said to be a *P-polynomial scheme* if the p-numbers can be represented in the form $p_i(k) = P_i(\xi_k)$ where $P_i(t)$ is a real polynomial of degree i in t and $\xi_0, \cdots, \xi_n$ are distinct real numbers. The orthogonality relations on the p-numbers show that $(P_i(t))_{i=0}^n$ is a system of *orthogonal polynomials.* There is a similar (dual) definition and a similar result for a *Q-polynomial scheme* (involving the q-numbers instead of the p-numbers) [37].

The P-polynomial property has a clear interpretation: R_i contains the pairs of points that are at distance i apart in the "generator graph" (X, R_1). In other words, (X, R_1) is a *distance-regular graph.* This subject was introduced by Biggs in 1969 (see [18]); it is treated in great detail by Brouwer, Cohen and Neumaier [26]. The dual notion of a Q-polynomial scheme is equally interesting and has been investigated by several authors [10], [54], [71], [72], [93], [94], [77], [81], [126]. It should be noted, however, that this notion does not have a simple "combinatorial meaning".

The theory of Q-polynomial schemes can be extended so as to include "continuous analogues" such as the *Euclidean sphere* and the *projective space* in a unified framework [48], [54], [65], [77], [81], [93], [116], [119]. In particular, the linear programming method can be applied to derive upper bounds for spherical codes with a given minimum distance and lower bounds for spherical designs with a given strength (see [45], [77], [81], [116], [135]).

If the point set X is endowed with the structure of an Abelian group and if the relations R_i are "translation invariant" with respect to that group, then the association scheme is said to be a *translation scheme* (with respect to the given group). This notion is equivalent to that of a commutative Schur ring, investigated in detail by Tamaschke [123]. There exists a *dual* translation scheme (with

[2] A related notion is the *outer distribution*, which enumerates the R_i-associates of each point $x \in X$ in the code Y.

respect to the dual group of X). If Y is an additive code in X, i.e., a subgroup of X, then there is a natural definition of an *annihilator code* Y° in the dual scheme. (The relation between Y and Y° generalizes the relation between a linear code Y in Hamming scheme and its orthogonal code $Y^\perp$.) The inner distributions of the codes Y and Y° are related by *generalized MacWilliams identities*, in the sense that they are the P-transform and Q-transform of each other (within scaling) [32], [37], [42].

Thus, the theory of translation schemes is quite interesting in that it provides the formal Krein duality with an actual duality interpretation. Furthermore, in this restricted context, there is a simple criterion to check whether a given additive code Y carries a "subscheme" of the translation scheme X, and to characterize the dual scheme of Y (see [37] in the case of the Hamming scheme).

This paper aims at giving a self-contained account of those parts of association scheme theory that are especially relevant to coding theory (in a wide sense), along the lines of the present introduction.

Section II contains the basic definitions; it is focused on the Bose–Mesner algebra and its formal duality. Section III introduces the subject of codes (and designs) in an association scheme, with special emphasis on the notions of the inner and outer distributions. This also includes the linear programming approach and a duality in bounding the sizes of codes and designs based on the existence of two orthogonality conditions. Section IV gives up to date bounds on fundamental parameters of codes and designs in P- and/or Q-polynomial schemes. Two extremum problems for systems of orthogonal polynomials are considered and their optimal solutions are used to describe the best known linear programming bounds. The results for Q-polynomials schemes are extended to the case of the unit Euclidean sphere. For P- and Q-polynomial schemes, three pairs of universal bounds and main asymptotic results are presented. Section V deals with translation schemes and their additive codes; it includes an introduction to $\mathbb{Z}_4$-additive binary codes.

II. Basic notions

A. Definitions and examples

Let X be a finite set of "points", with $|X| \geq 2$. For an integer $n \geq 1$, consider a set $R = \{R_0, R_1, \cdots, R_n\}$ of $n+1$ nonempty *binary relations* R_i on X (i.e., $R_i \subset X^2$), forming a *partition* of the Cartesian square X^2 of X.

For integers a and b with $a \leq b$, we shall use the notation N_b^a for the integer interval $\{a, a+1, \cdots, b\}$ and put $N_n := N_n^0 = \{0, 1, \cdots, n\}$.

Definition 1: The pair (X, R) is said to be an *n-class association scheme* if

(a) R_0 is the diagonal, i.e., $R_0 = \{(x,x) \in X^2 : x \in X\}$.

(b) For $i \in N_n$, the converse $R_i^\cup := \{(x,y) \in X^2 : (y,x) \in R_i\}$ of R_i belongs to R.

(c) There exist integer numbers $p_{i,j}^k$, called *intersection numbers*, with $p_{i,j}^k = p_{j,i}^k$, such that, for each pair $(x,y) \in R_k$, the number of points $z \in X$ with $(x,z) \in R_i$ and $(z,y) \in R_j$ is equal to the constant $p_{i,j}^k$ (for $i,j,k \in N_n$).

Condition (b) induces a *pairing* $i \mapsto i^\sigma$ over N_n, defined by $R_{i^\sigma} = R_i^\cup$. The number p_{i,i^σ}^0 is denoted by v_i and is called the *valency* (or the degree) of the directed graph (X, R_i); it counts the points $z \in X$ with $(x,z) \in R_i$, for any fixed $x \in X$. Clearly, $v_{i^\sigma} = v_i$ and $\sum_{i=0}^n v_i = |X|$.

In most coding theoretic applications, the definition above can be made more restrictive. The association scheme (X,R) is said to be *symmetric* if all its relations R_i are symmetric. Thus, condition (b) is replaced by

(b)* $R_i = R_i^\cup$, for each $i \in N_n$.

In other words, a symmetric association scheme has a *trivial pairing*, $i = i^\sigma$ for all i. (Notice that the identity $p_{i,j}^k = p_{j,i}^k$ in (c) can be omitted in the symmetric case, since it becomes a consequence of the other requirements.)

In particular, a 2-class association scheme ($n = 2$) is equivalent to a *strongly regular graph* [20], [105] in the symmetric case, and to a *skew conference matrix* in the non-symmetric case [13], [56].

Note that we can consider an association scheme (X,R) with an ordering of R_i $i = 0, 1, ..., n$, as a space X with the function $\partial_R : X^2 \to N_n$ which is defined as follows:

$$\partial_R(x,y) = i \text{ if and only if } (x,y) \in R_i. \quad (1)$$

In the symmetric case this function has, in particular, the properties $\partial_R(x,y) = 0$ if and only if $x = y$ and $\partial_R(x,y) = \partial_R(y,x)$, but, in general, does not satisfy the triangle inequality and hence is not a distance function. On the other hand, a metric space X with a distance function $\partial(x,y)$ which takes values from N_n is a symmetric n-class association scheme (X,R) with $\partial_R(x,y) = \partial(x,y)$ if and only if for any $i,j \in N_n$ and $x,y \in X$, the number

$$\lambda_{i,j}(x,y) := |\{z \in X : \partial(x,z) = i, \partial(z,y) = j\}| \quad (2)$$

depends only on i, j, and $\partial(x,y)$. In fact we state that this is true for the first two examples below.

Example 1 Let $X = \mathbf{F}^n$ be the nth Cartesian power of a finite *alphabet* $\mathbf{F}$, with $|\mathbf{F}| = q \geq 2$. Let $\partial_H : X^2 \to N_n$ denote the *Hamming distance function*:

$$\partial_H(x,y) := |\{j \in N_n^1 : x_j \neq y_j\}|.$$

Then (X,R) with $\partial_R(x,y) = \partial_H(x,y)$ is a symmetric n-class association scheme, called the *Hamming scheme* and denoted by H_q^n. It appears as the natural framework of the classical theory of "block codes" [82], [88]. When q is a prime power, $\mathbf{F}$ can be endowed with the structure of the *finite field* $\mathbb{F}_q$. In this case, $\partial_H(x,y) = w_H(x-y)$, where $w_H : X \to N_n$ is the *Hamming weight function*, given by $w_H(x) := |\{j \in N_n^1 : x_j \neq 0\}|$. More generally, this applies to the case where $\mathbf{F}$ has the structure of an *additive Abelian group*. (No multiplicative operation is required here.)

Example 2 Let X be the set of binary v-tuples of a fixed weight n, with $1 \leq n \leq \lfloor v/2 \rfloor$. Thus, $X := \{x \in \{0,1\}^v : w_H(x) = n\}$. For $i \in N_n$, define $R_i := \{(x,y) \in X^2 :$

$\partial_H(x,y) = 2i\}$, and $R := \{R_0, R_1, \cdots, R_n\}$. Then (X, R) is a symmetric n-class association scheme, called the *Johnson scheme* and denoted by J_n^v. It is a "subscheme" of the binary Hamming scheme H_2^v with $\partial_R(x,y) = \frac{1}{2}\partial_H(x,y)$. The Johnson scheme plays a useful role in combinatorial coding theory.

Example 3 Let $\mathbf{F} = \{\alpha_0, \alpha_1, \cdots, \alpha_{q-1}\}$, and $X := \mathbf{F}^v$. The *composition* of a point x in X is the integer q-tuple $(s_0(x), s_1(x), \cdots, s_{q-1}(x))$ defined by $s_l(x) := |\{j \in N_v^1 : x_j = \alpha_l\}|$. Assume that $\mathbf{F}$ is an Abelian group. Define a set R of binary relations R_i on X as follows. A pair (x, y) in X^2 belongs to a certain relation R_i if and only if the difference $x - y$ has a specified composition. Then (X, R) is an n-class association scheme, with $n = \binom{v+q-1}{q-1} - 1$, called the *composition scheme*. It is symmetric when $x = -x$ for all $x \in X$, i.e., when $\mathbf{F}$ is an elementary Abelian 2-group (of order $q = 2^m$). In particular, the composition scheme with $q = 2$ reduces to the binary Hamming scheme H_2^n.

There are several other families of association schemes that have interesting applications in coding theory. Let us mention five of them: (i) the association scheme relative to the *split weight enumerator* [42], [88]; (ii) the *Lee scheme* [124]; (iii) the *non-binary Johnson scheme* [1], [124], [125]; (iv) the association scheme of *$m \times n$ matrices over a finite field* [41] (which has applications in crisscross error correcting codes [52], [102]); (v) the association scheme of *$n \times n$ skew-symmetric matrices over a finite field* [43]. In the last two cases, the relations R_i are defined from the *rank metric* over the matrix set X.

Of course, there exist applications of association schemes outside the area of coding theory (in a wide sense). It is especially worth saying that association schemes have recently found considerable interest in *spin model theory* (a branch of mechanical statistics). The idea is due to Jaeger (see [63] and the references therein).

Finally, let us mention some constructions that produce association schemes from other association schemes. In particular, there is the notion of the *extension* [37], *product* [54], and *merging* [32] of association schemes.

B. The Bose–Mesner algebra

It proves very useful to investigate a combinatorial structure such as an association scheme (X, R) by matrix algebra methods. Let $\mathbb{C}(X^2)$ denote the set of square complex matrices M of order $|X|$, where rows and columns are labeled with the points $x \in X$, and the (x, y) entry of M is denoted by $M(x, y)$. The directed graph (X, R_i) is represented by its *adjacency matrix* $D_i \in \mathbb{C}(X^2)$, defined by

$$D_i(x,y) := \begin{cases} 1 & \text{for } (x,y) \in R_i \\ 0 & \text{for } (x,y) \notin R_i \end{cases} \tag{3}$$

Definition 2: Let (X, R) be an n-class association scheme (see Definition 1). The *Bose–Mesner algebra* of (X, R), denoted by $\mathcal{A}$, is the complex vector space generated by the adjacency matrices D_i, that is

$$\mathcal{A} := \{c_0 D_0 + c_1 D_1 + \cdots + c_n D_n \,:\, c_0, c_1, \cdots, c_n \in \mathbb{C}\}. \tag{4}$$

From the fact that R is a partition of X^2 and from condition (a), it follows that $\mathcal{A}$ *contains the all-one matrix J and the unit matrix I*, since

$$D_0 + D_1 + \cdots + D_n = J, \quad D_0 = I. \tag{5}$$

Condition (b) says that $\mathcal{A}$ is *closed under conjugation* ($M \mapsto \overline{M}$) *and under transposition* ($M \mapsto M^T$), whence under conjugate transposition ($M \mapsto M^* := \overline{M}^T$), since

$$\overline{D_i} = D_i, \quad D_i^T = D_{i^\sigma}. \tag{6}$$

Condition (c) says that $\mathcal{A}$ is *closed under matrix multiplication*, and that multiplication in $\mathcal{A}$ is *commutative*, since

$$D_i D_j = \sum_{k=0}^{n} p_{i,j}^k D_k = D_j D_i. \tag{7}$$

This shows that the $(n+1)$-dimensional vector space $\mathcal{A}$ defined by (4) has the structure of a *commutative algebra* (over $\mathbb{C}$). As indicated in Definition 2 (with some anticipation in the use of the term "algebra"), $\mathcal{A}$ is usually referred to as the *Bose–Mesner algebra* (or *adjacency algebra*) of the association scheme (X, R) (see [21]).

For a symmetric association scheme, we have $D_i^T = D_i$ for all i. In this case, we can define the Bose–Mesner algebra over the *reals*, i.e., replace $\mathbb{C}$ by $\mathbb{R}$ in (4).

The adjacency algebra $\mathcal{A}$ is known to be *semi-simple*. This means that there exists a unitary matrix U of order $|X|$ that reduces each matrix $M \in \mathcal{A}$ to a diagonal form $\Delta_M = U^{-1}MU$. As a consequence, $\mathcal{A}$ possesses a unique basis of *irreducible idempotent matrices* $E_0, E_1, \cdots, E_n$, which are mutually orthogonal:

$$E_k E_l = \delta_{k,l} E_k, \quad \text{for } k, l \in N_n. \tag{8}$$

In particular, $E_0 = |X|^{-1}J$. The rank of E_k will be denoted by m_k, and the numbers $m_0, m_1, \cdots, m_n$ will be referred to as the *multiplicities* of the association scheme (X, R). By definition, E_k has eigenvalues 1 and 0 with multiplicities m_k and $|X| - m_k$. Notice that $m_0 = 1$ and $\sum_{k=0}^n m_k = |X|$. Considering the inner product

$$\langle v, w \rangle = \frac{1}{|X|} \sum_{x \in X} v(x)\overline{w(x)}, \tag{9}$$

for complex functions v, w defined on X, one can represent the matrix E_k in the form

$$E_k(x,y) = \frac{1}{|X|} \sum_{j=1}^{m_k} v_{k,j}(x)\overline{v_{k,j}(y)} \tag{10}$$

where $\{v_{k,j} : j = 1, ..., m_k\}$ is an arbitrary orthonormal basis of the linear space V_k generated by the columns of E_k. (It is orthonormal with respect to (9).)

Definition 3: The *p-numbers* of an n-class association scheme (X, R) are the complex numbers $p_i(k)$, with $k, i \in N_n$, defined from the expansion of the adjacency matrices

D_i in the basis of the irreducible idempotent matrices E_k of the algebra $\mathcal{A}$, i.e.,

$$D_i = \sum_{k=0}^{n} p_i(k) E_k, \quad \text{for } i \in N_n. \tag{11}$$

Analogously, the *q-numbers* of (X, R) are the complex numbers $q_k(i)$, with $i, k \in N_n$, defined from the inverse expansion, within the normalizing factor $|X|$, i.e.,

$$|X| E_k = \sum_{i=0}^{n} q_k(i) D_i, \quad \text{for } k \in N_n. \tag{12}$$

These numbers play a major role in the theory. It follows from (11) that the p-number $p_i(k)$ is the *eigenvalue* of D_i relative to the m_k-dimensional space V_k spanned by the columns of E_k. In particular, $p_i(0) = v_i$ (valency of R_i). Notice that $q_k(0) = m_k$ (rank of E_k). In view to (3) and (1), (12) can be written in the form

$$|X| E_k(x, y) = q_k(\partial_R(x, y)). \tag{13}$$

If the association scheme (X, R) is *symmetric*, then its p-numbers and q-numbers are *real.*

Let $F(N_n)$ denote the linear space of complex (or real in the symmetric case) functions defined on N_n. In particular, the p-numbers $p_i(k)$ and the q-numbers $q_k(i)$ are values of the *p-functions* $p_i \in F(N_n)$ and *q-functions* $q_k \in F(N_n)$ which forms two bases of $F(N_n)$. This implies that any function $h \in F(N_n)$ has a unique expansion over either of these bases:

$$h = \sum_{i=0}^{n} h_i(p) p_i, \quad h = \sum_{k=0}^{n} h_k(q) q_k. \tag{14}$$

The following result expresses the well-known *orthogonality relations* for the p-functions and q-functions. It appears as a consequence of (8), basically.

Theorem 1: The *p-functions* $p_i \in F(N_n)$ $(i = 0, 1, ..., n)$ are pairwise orthogonal on N_n with respect to the multiplicities m_k and the *q-functions* $q_k \in F(N_n)$ $(k = 0, 1, ..., n)$ are pairwise orthogonal on N_n with respect to the valencies v_i. More precisely,

$$\sum_{k=0}^{n} m_k \, \overline{p_i(k)} \, p_j(k) = |X| v_i \, \delta_{i,j},$$
$$\sum_{i=0}^{n} v_i \, \overline{q_k(i)} \, q_l(i) = |X| m_k \, \delta_{k,l}.$$

In view of the fact that the relations (11) and (12) are inverse of each other, the p-numbers and the q-numbers are related by

$$m_k \overline{p_i(k)} = v_i q_k(i). \tag{15}$$

Example 1 (continued) For given values of n (the "length") and q (the "alphabet size"), and for $k \in N_n$, we define the *Krawtchouk polynomial* $K_k^n(z)$ as follows [68]:

$$K_k^n(z) := \sum_{j=0}^{k} (-1)^j (q-1)^{k-j} \binom{z}{j} \binom{n-z}{k-j}. \tag{16}$$

Clearly, $K_k^n(z)$ is a polynomial of degree k in z. The p-numbers and the q-numbers of H_q^n are the values assumed by the Krawtchouk polynomials at the integer points $0, 1, \cdots, n$. More precisely,

$$p_i(k) = K_i^n(k), \quad q_k(i) = K_k^n(i). \tag{17}$$

The valencies and multiplicities are $v_k = m_k = \binom{n}{k}(q-1)^k$.

Example 2 (continued) For given values of v (the "length") and n (the "weight"), the valencies and multiplicities of the Johnson scheme J_n^v are given by

$$v_i = \binom{n}{i}\binom{v-n}{i}, \quad m_k = \binom{v}{k} - \binom{v}{k-1}.$$

For $k, i \in N_n$, we define the *Hahn polynomial*[3] $H_k(z)$ and the *dual Hahn polynomial* $\tilde{H}_i(z)$ as follows [66]:

$$H_k(z) = m_k \sum_{j=0}^{k} (-1)^j \frac{\binom{k}{j}\binom{v+1-k}{j}}{\binom{n}{j}\binom{v-n}{j}} \binom{z}{j} \tag{18}$$

$$\tilde{H}_i(z) = \sum_{j=0}^{i} (-1)^{i-j} \binom{n-j}{i-j} \binom{n-z}{j} \binom{v-n+j-z}{j} \tag{19}$$

Clearly, $H_k(z)$ is a polynomial of degree k in z. It is easily seen that $\tilde{H}_i(z)$ is a polynomial of degree i in $z(v+1-z)$. The p-numbers and the q-numbers of J_n^v can be determined (see [37]) from these polynomials by

$$p_i(k) = \tilde{H}_i(k), \quad q_k(i) = H_k(i). \tag{20}$$

C. Formal duality

The adjacency matrices D_i and the idempotent matrices E_k play *dual roles* in the theory. This formal duality, which interchanges the p-numbers and the q-numbers, will be referred to as the *Krein duality* [8]. Let us examine this subject in some detail. The Bose–Mesner algebra $\mathcal{A}$ is closed not only under ordinary matrix multiplication $(M, N) \mapsto MN$, but also under *pointwise* (or *Hadamard*) *multiplication* $(M, N) \mapsto M \circ N$, defined by $(M \circ N)(x, y) := M(x, y) N(x, y)$. This stems from the fact that the adjacency matrices D_i are "idempotent" and "mutually orthogonal" with respect to the pointwise product:

$$D_i \circ D_j = \delta_{i,j} D_i, \quad \text{for } i, j \in N_n. \tag{21}$$

The formal Krein duality under discussion *permutes the roles of the matrix product and the pointwise product.* Thus, the identities (8) and (21) are dual of each other. As duals of (5) and (6), we have

$$E_0 + E_1 + \cdots + E_n = I, \quad E_0 = |X|^{-1} J,$$
$$E_k^* = E_k, \quad E_k^T = E_{k^\tau},$$

where $k \mapsto k^\tau$ is a well-defined pairing over N_n. For a symmetric association scheme, this pairing is trivial: $E_k^T = E_k$ for all k.

[3]In fact, these are the Hahn polynomials of *spherical type.*

Let us stress the fact that the idempotent matrices E_k are *Hermitian* and *nonnegative definite*, since their eigenvalues are 0 and 1. (It can be viewed as a dual of the property of adjacency matrices, having entries 0 and 1.)

The following property is essential for the linear programming method introduced in Section III below.

Theorem 2: For any function $h \in F(N_n)$ the matrix $h(\partial_R(x,y))$ is Hermitian and nonnegative definite if and only if $h_k(q) \geq 0, k = 0, 1, ..., n$.

Next, we examine the dual of identity (7), that is

$$|X|(E_k \circ E_l) = \sum_{m=0}^{n} q_{k,l}^m E_m. \tag{22}$$

The numbers $q_{k,l}^m$ defined from (22) are usually called the *Krein parameters*; they are the duals of the intersection numbers $p_{i,j}^k$. In particular, $q_{k,k^\tau}^0 = m_k$, and $q_{k,l}^0 = 0$ when $l \neq k^\tau$. Thus, the multiplicities m_k are the duals of the valencies v_i. Notice that $E_k \circ E_l$ is a Hermitian nonnegative definite matrix according to (10). Hence, by (13) and Theorem 2 the Krein parameters satisfy $q_{k,l}^m \geq 0$ (see [55], [101]).

By use of (7) and (22) we deduce that the intersection numbers $p_{i,j}^s$ and the Krein parameters $q_{k,l}^t$ are the *linearization factors* relative to the p-numbers $p_i(k)$ and to the q-numbers $q_k(i)$, respectively, in the sense that

$$p_i(k)p_j(k) = \sum_{s=0}^{n} p_{i,j}^s p_s(k), \tag{23}$$

$$q_k(i)q_l(i) = \sum_{m=0}^{n} q_{k,l}^m q_m(i). \tag{24}$$

Two n-class association schemes (X,R) and (X',R'), with $|X| = |X'|$, are *formal dual* of each other if the p-numbers of (X,R) are the q-numbers of (X',R'), and conversely, i.e.,

$$p_k'(i) = q_k(i), \quad q_i'(k) = p_i(k). \tag{25}$$

A necessary condition for an association scheme to have a formal dual is that its Krein parameters be integers (since they are the intersection numbers of the dual).

Example 1 (continued) In view of (17), the Hamming scheme H_q^n is formally self-dual. In fact, it is actually self-dual in the strong sense of "duality in translation schemes" (see Section V below).

Example 2 (continued) The Johnson scheme J_n^v has no formal dual. However, the general Krein duality applies; it permutes the Hahn and dual Hahn polynomials.

D. The group case, and generalizations

In Examples 1, 2, 3 of Section II-A, the *regularity properties* defining the association scheme structure are induced by some *symmetry properties*, i.e., by a certain "group of automorphisms". We now say a few words on this subject (see [10], [58], [132]). Let G be a transitive permutation group acting on the point set X. It induces a partition of X^2 into a well-defined set $R = \{R_0, R_1, \cdots, R_n\}$ of *orbits* R_i. (By definition, such an orbit R_i contains the images $(x_i^g, y_i^g) \in X^2$ of a fixed pair $(x_i, y_i) \in X^2$ under all mappings $g \in G$.) The resulting structure (X,R) satisfies the conditions (a), (b), (c) of an association scheme, except possibly the "commutativity condition" $p_{i,j}^k = p_{j,i}^k$. In any case, the adjacency matrices D_i form the basis of a subalgebra of $\mathbb{C}(X^2)$. (It is called the *Hecke algebra*.) This algebra is commutative if and only if $p_{i,j}^k = p_{j,i}^k$ (for all i, j, k).

Example 1 (continued) Let G be the permutation group on $X := \mathbf{F}^n$ generated by two types of mappings:

(i) a permutation on the n coordinates;

(ii) in each coordinate position, a permutation on the q alphabet symbols.

This group has order $n!(q!)^n$, and it is transitive on X. The corresponding structure (X,R) is the Hamming scheme H_q^n. In particular, the binary Hamming scheme H_2^n arises from the (complete) *monomial group* $G = M_n$ of degree n, containing the matrices of order n that have one nonzero element, equal to ± 1, in each row and each column.

Example 2 (continued) Let G be the *symmetric group* S_v of degree v, containing the $v!$ permutations on v coordinates. It acts in a natural way on the set X of binary v-tuples of weight n. The corresponding structure (X,R) is the Johnson scheme J_n^v.

The notion of an association scheme (X,R) can be generalized, by omitting the commutativity requirement $p_{i,j}^k = p_{j,i}^k$. A further extension is obtained by relaxing the "homogeneity condition" (a) in Definition 1. In general, it is only required that the diagonal relation $\{(x,x) : x \in X\}$ be a *union* of some relations belonging to the set R. Thus we arrive at a combinatorial structure (X,R) called a *coherent configuration* [59] (equivalent to a *cellular ring* [50]). The group theoretic counterpart of this general structure is obtained by leaving out the transitivity assumption.

Certain infinite metric spaces occur as analogues of association schemes that are important in coding theoretic applications. We call *distance-transitive* (or two-point homogeneous [130]) a connected compact metric space X with the distance function $\partial(\cdot,\cdot)$ and the isometry group G, if for any $x_1, x_2, y_1, y_2 \in X$ the equality $\partial(x_1,y_1) = \partial(x_2,y_2)$ implies the existence of some $g \in G$ such that $x_2 = x_1^g$ and $y_2 = y_1^g$. As an example of a distance-transitive space we mention the unit Euclidean sphere S^{n-1} in $\mathbb{R}^n$ (considered in more detail in Section IV-E) whose isometry group consists of all orthogonal matrices of order n. A distance-transitive space X has many strong properties [9], [53], [65], [120], [129]. The isometry group G of X acts *transitively* on X and hence there exists a unique *normalized invariant measure* μ ($\mu(X) = 1$ and $\mu(A^g) = \mu(A)$ for any measurable $A \subseteq X$ and $g \in G$). If n_X is the *diameter* of X, then for any (real) $i, j \in [0, n_X]$ and $x, y \in X$, $\mu\{z \in X : \partial(x,z) \leq i, \partial(z,y) \leq j\}$ (cf. (2)) depends only on i, j, and $\partial(x,y)$. For any *invariant* function $H(x,y)$ on X^2 (this means that $H(x^g, y^g) = H(x,y)$ for any $g \in G$) there exists a function h on $[0, n_X]$ such that $H(x,y) = h(\partial(x,y))$. Continuous invariant functions $H(x,y)$ on X^2 form a commutative algebra $\mathcal{A}$ with respect to the operations of addition and *convolution*:

$$F * H(x,z) = \int_X F(x,y)H(y,z)d\mu(y).$$

In the linear space V of continuous functions $v(x)$ on X with the inner product

$$\langle v, w\rangle = \int_X v(x)\overline{w(x)}d\mu(x) \tag{26}$$

(cf. (9)), the unitary representation $L(g)$ of G defined as follows: $L(g)v(x) = v(x^{g^{-1}})$, decomposes into a countable direct sum of pairwise inequivalent irreducible representations $L_k(g)$ acting on (mutually orthogonal) subspaces V_k, $k = 0, 1, \ldots$, of continuous functions. Each subspace V_k has a finite dimension m_k (V_0 consists of constants and $m_0 = 1$) and is *invariant* (i.e., if $v(x) \in V_k$, then $v(x^g) \in V_k$ for any $g \in G$). The invariant functions (cf. (10))

$$\tilde{E}_k(x,y) = \sum_{j=1}^{m_k} v_{k,j}(x)\overline{v_{k,j}(y)}, \quad k = 0, 1, \ldots \tag{27}$$

where $\{v_{k,j}(x) : j = 1, \ldots, m_k\}$ is an arbitrary orthonormal (with respect to (26)) basis of V_k, form a basis of $\mathcal{A}$ consisting of irreducible idempotents, which are mutually orthogonal: $\tilde{E}_k * \tilde{E}_l = \delta_{k,l}\tilde{E}_k$, $k, l = 0, 1, \ldots$. The corresponding "q-functions" q_k on $[0, n_X]$ such that $\tilde{E}_k(x,y) = q_k(\partial(x,y))$ are real and satisfy the following orthogonality and normalization conditions:

$$\int_0^{n_X} q_k(z)q_l(z)d\tilde{\mu}(z) = m_k\delta_{k,l}, \quad q_k(0) = m_k, \tag{28}$$

where $\tilde{\mu}$ is the measure on $[0, n_X]$ such that $\tilde{\mu}(A) = \mu\{y \in X : \partial(x_0, y) \in A\}$ (this does not depend on $x_0 \in X$). For any element $H(x,y) = h(\partial(x,y))$ of $\mathcal{A}$, the series $\sum_{k=0}^{\infty} h_k q_k(z)$ with $h_k = (m_k)^{-1}\int_0^{n_X} h(z)q_k(z)d\tilde{\mu}(z)$ converges to $h(z)$ on $[0, n_X]$. Moreover, for these functions q_k, $k = 0, 1, \ldots$, an analog of Theorem 2 is valid and the linearization factors $q_{k,l}^m$ are nonnegative. All (infinite) compact distance transitive spaces have been classified in [130] as the unit Euclidean spheres S^{n-1}, the projective spaces in n dimensions over $\mathbb{R}, \mathbb{C}$, and quaternions $\mathbb{H}$ ($n = 2, 3, \ldots$), and the Cayley projective plane.

Note that the definition and all properties of distance-transitive spaces are also correct for finite metric spaces, and any finite distance-transitive metric space X is an n-class symmetric association scheme (X, R) with $\partial_R(x,y) = \partial(x,y)$ and n equal to the number of nonzero values of $\partial(x,y)$. In particular, the Hamming and Jonhson spaces are distance-transitive. The fact that in the case of finite spaces $\tilde{E}_k(x,y) = |X|E_k(x,y)$ (compare (10) with (27)) is explained by the distinctness between the product and convolution of matrices.

III. Codes and designs

In coding theory and related subjects, an association scheme (such as the Hamming scheme) should mainly be viewed as a "structured space" in which the objects of interest (such as codes, or designs) are living.

Let Y be a nonempty subset of the point set X of an association scheme (X, R). Then Y will be called a *code in* (X, R). (In certain contexts, Y is preferably called a *design.*) We now introduce the important concept of the inner distribution of a code.

A. Inner distribution

Definition 4: The *inner distribution* of a code Y in an n-class association scheme (X, R) is the rational $(n+1)$-tuple $(a_0, a_1, \cdots, a_n)$ where $|Y|a_i$ counts the pairs of points in Y that belong to the relation R_i. Formally,

$$a_i = a_i(Y) := \frac{1}{|Y|}|Y^2 \cap R_i|, \quad \text{for } i \in N_n. \tag{29}$$

A code Y in the Hamming scheme H_q^n is nothing but a *q-ary code of length n.* The inner distribution of Y is its *(Hamming) distance distribution.* In effect, $|Y|a_i$ counts the pairs of codewords $y, y' \in Y$ with $\partial_H(y, y') = i$.

Coding theorists are often interested in a code having a specified set of admissible distances (in particular: a specified minimum distance). In the general framework of association schemes, this notion extends as follows.

Definition 5: Let D be a subset of N_n^1. A code Y in (X, R) is called a *D-code* if all pairs of distinct points in Y belong to the admissible relation $\bigcup_{i\in D} R_i$. In terms of the inner distribution, this becomes $a_i(Y) = 0$ for each $i \in N_n^1 \setminus D$.

Consider, for a while, the familiar situation where Y is a *linear code of length n over the field* $\mathbb{F}_q$ (in the Hamming scheme). Then the distance distribution of Y reduces to its *weight distribution*: $a_i(Y)$ counts the codewords $y \in Y$ with $w_H(y) = i$. From the linear code Y we can define its *orthogonal code* (often called the dual code), that is $Y^\perp := \{x \in \mathbb{F}_q^n : xy^T = 0 \text{ for all } y \in Y\}$. The weight distributions of Y and $Y^\perp$ are related by the *MacWilliams identities* [86], [87]. These are well-defined linear relations involving the Krawtchouk polynomials (16). (We shall go back to this subject in Section V.)

As a result, the "Krawtchouk-MacWilliams transform" of the distance (or weight) distribution of a linear code Y yields *nonnegative real numbers*, which can be interpreted as the components $a_k(Y^\perp)$ of the distance distribution of the orthogonal code $Y^\perp$. It turns out that this nonnegativity result can be extended to *arbitrary codes* (in a Hamming scheme), even though the orthogonal code notion is lost. Moreover, as shown below, the result extends to codes in *any* association scheme.

Definition 6: Let $q_k(i)$ denote the q-numbers of an n-class association scheme (with $i, k \in N_n$). The *Q-transform* of a complex $(n+1)$-tuple $(a_i)_{i=0}^n$ is the complex $(n+1)$-tuple $(a'_k)_{k=0}^n$ given by

$$a'_k := \sum_{i=0}^{n} a_i q_k(i), \quad \text{for } k \in N_n. \tag{30}$$

Note that this definition of the Q-transform in fact depends on the choice of an ordering of the functions q_k (or the matrices E_k), $k=0,1,...,n$. From Theorem 1 and (15) it follows that

$$|X|a_i = \sum_{k=0}^{n} a'_k p_i(k). \tag{31}$$

Moreover (see (14)), for any $h \in F(N_n)$,

$$\sum_{i=0}^{n} a_i h(i) = \sum_{k=0}^{n} a'_k h_k(q), \tag{32}$$

$$\sum_{k=0}^{n} a'_k h(k) = |X| \sum_{i=0}^{n} a_i h_i(p). \tag{33}$$

Theorem 3: (Generalized MacWilliams inequalities [37]) Let $(a_i(Y))_{i=0}^n$ be the inner distribution of a code Y in (X,R), and let $(a'_k(Y))_{k=0}^n$ be its Q-transform. Then $a'_k(Y) \geq 0$ (i.e., $a'_k(Y) \in \mathbb{R}_+$) for each $k \in N_n$.

The proof is quite easy since, in view of (13) and (10),

$$|Y|a'_k(Y) = \sum_{x,y\in Y} q_k(\partial_R(x,y)) = \sum_{j=1}^{m_k} \left| \sum_{x \in Y} v_{k,j}(x) \right|^2 \geq 0. \tag{34}$$

This also shows that $a'_k(Y)$ is an *averaging* parameter of a code Y. In this connection note that $a'_k(X)=0$ for all $k=1,...,n$.

The Q-transform of the inner distribution of a code Y will sometimes be referred to as the "dual (inner) distribution" of Y.

Definition 7: Let D be a subset of N_n^1. A code Y in (X,R) is called a *D-design* if the Q-transform of its inner distribution satisfies $a'_k(Y)=0$ for each $k \in N_n^1 \setminus D$.

Example 1 (continued) In H_q^n, a $N_n^{\tau+1}$-design Y is an *orthogonal array of strength* τ (see [37], [40]). This means that the restriction of Y to any set of τ coordinates shows all τ-tuples of alphabet symbols appearing the same number of times [98]. Orthogonal arrays are closely related to "resilient functions" and to "correlation immune functions" which occur in some cryptography applications [16], [33], [79], [113], [122].

Example 2 (continued) In J_n^v, a $N_n^{\tau+1}$-design Y is a *combinatorial τ-design* (see [37]). This is a collection of blocks of size n, out of a point set of size v, such that all τ-subsets of the v-set are contained in the same number of blocks [15], [62]. There is a close connection between coding theory and design theory [3], [5]. It is interesting to point out that a 3-parameter class of Hahn polynomials (larger than the "spherical class" involved in the q-numbers) plays a significant role in the theory of combinatorial τ-designs [133], [134] and in an extension thereof [28].

B. The linear programming method

Theorem 3 strongly suggests using linear programming to find bounds on the size of a code Y characterized by some linear constraints on its inner distribution $(a_i(Y))_{i=0}^n$. In particular, this method leads to upper bounds for D-codes and to lower bounds for D-designs. We shall use the "non-standard forms" of the linear programming problem.

For simplicity we assume, in this subsection, that (X,R) is a *symmetric* association scheme, which implies that the p- and q-numbers are *real*. For the problems that we are considering here, this entails no loss of generality. While simultaneously considering p-functions and q-functions it is convenient to use the letter u instead of either p or q, and use $\overline{u}$ for the other one. For any $h \in F(N_n)$ with the expansion $h=\sum_{k=0}^n h_k(u)u_k$ we put

$$\Omega_u(h) = h(0)/h_0(u) \text{ if } h_0(u) \neq 0.$$

For any $D \subseteq N_n^1$, we say that $h \in F(N_n)$ has the property $\mathfrak{A}_u(D)$ if

$$h_0(u) > 0, \quad h_i(u) \geq 0 \text{ for } i \in N_n^1,$$

$$h(0) > 0, \quad h(i) \leq 0 \text{ for } i \in D,$$

and has the property $\mathfrak{B}_u(D)$ if

$$h_0(u) > 0, \quad h_i(u) \leq 0 \text{ for } i \in D,$$

$$h(0) > 0, \quad h(i) \geq 0 \text{ for } i \in N_n^1.$$

Let

$$A_u(X,D) = \min \Omega_u(h),$$

where the minimum is taken over all functions $h \in F(N_n)$ with the property $\mathfrak{A}_u(D)$ and

$$B_u(X,D) = \max \Omega_u(h)$$

where the maximum is taken over all functions $h \in F(N_n)$ with the property $\mathfrak{B}_u(D)$. It should be noted that both extremum problems are linear programming problems, because without loss of generality one can assume that $h_0(u)=1$ and then $\Omega_u(h)=h(0)=\sum_{j=0}^n h_j(u)u_j(0)$.

The following two results are obtained respectively with the help of (32) and (33) with $a_i=a_i(Y)$ and $a'_i=a'_i(Y)$. One also makes use of $a_0(Y)=1, a'_0(Y)=\sum_{i=0}^n a_i(Y) = |Y|, a_i(Y) \geq 0, a'_i(Y) \geq 0$ (by Theorem 3), $a_i(Y)=0$ if $i \in N_n^1 \backslash D$ when Y is a D-code, and $a'_i(Y)=0$ if $i \in N_n^1 \backslash D$ when Y is a D-design.

Theorem 4: [37] If Y is a D-code and $h \in F(N_n)$ has the property $\mathfrak{A}_q(D)$, then

$$|Y| \leq A_q(X,D) \leq \Omega_q(h). \tag{35}$$

If Y is a D-design and $h \in F(N_n)$ has the property $\mathfrak{B}_q(D)$, then

$$|Y| \geq B_q(X,D) \geq \Omega_q(h). \tag{36}$$

In each case, equality $|Y|=\Omega_q(h)$ holds if and only if

$$a_i(Y)h(i) = a'_i(Y)h_i(q) = 0, \ i \in N_n^1.$$

Theorem 5: [37] If Y is a D-code and $h \in F(N_n)$ has the property $\mathfrak{B}_p(D)$, then

$$|Y| \leq |X|/B_p(X,D) \leq |X|/\Omega_p(h). \tag{37}$$

If Y is a D-design and $h \in F(N_n)$ has the property $\mathfrak{A}_p(D)$, then

$$|Y| \geq |X|/A_p(X,D) \geq |X|/\Omega_p(h). \tag{38}$$

In each case, equality $|Y| = |X|/\Omega_p(h)$ holds if and only if

$$a_i(Y)h_i(p) = a'_i(Y)h(i) = 0, \quad i \in N_n^1.$$

The necessary and sufficient conditions for these bounds to be sharp have many useful consequences, a nice example being the (generalized) *Lloyd condition* for perfect codes [17], [37], [70], [84], [101], [117]. Note that Theorems 4 and 5 imply that the functions h for which $|Y| = \Omega_q(h)$ or $|Y| = |X|/\Omega_p(h)$ holds are *optimal* solutions of the corresponding extremum problems.

Coding theorists are especially interested in applying Theorems 4 and 5 to the class of codes with a *specified minimum distance* d, which are N_n^d-codes in H_q^n and J_n^v. It was shown in [35] that the classical "elementary bounds" such as the *Hamming*, *Plotkin*, and *Singleton* bounds occur as simple cases of these theorems. In the next section we give bounds which are obtained with the help of *optimal solutions of some extremum problems for systems of orthogonal polynomials.* Combinatorial proofs of some of these bounds are unknown. It should be noted that the bounds of Theorems 4 and 5 can be improved by the same linear programming method if one knows an additional information about $a_i(Y)$ and $a'_i(Y)$, $i \in N_n$ (not only their nonnegativity). It was successfully used in the analysis of concrete codes (see [14], [27], [88]).

In conclusion of this subsection we verify that there exists a duality in bounding the sizes of D-codes and D-designs [78], [80]. For any $h \in F(N_n)$ and u (which is again either p or q), we define an u-dual function $h^{(u)}$ to h as follows:

$$h^{(u)} := |X|^{-\frac{1}{2}} \sum_{i=0}^{n} h(i) u_i. \tag{39}$$

Using Theorem 1 and (15) one can show that $h^{(u)}(i) = |X|^{1/2} h_i(\overline{u})$ and hence $h = (h^{(u)})^{(\overline{u})}, \Omega_{\overline{u}}(h)\Omega_u(h^{(u)}) = |X|$, and h has the property $\mathfrak{A}_{\overline{u}}(D)$ or $\mathfrak{B}_{\overline{u}}(D)$ if and only if $h^{(u)}$ has respectively the property $\mathfrak{B}_u(D)$ or $\mathfrak{A}_u(D)$. In particular, this implies the equivalence of the bounds (35) and (37) and also (36) and (38).

Theorem 6: [80] For any symmetric association scheme X and any $D \subseteq N_n^1$,

$$A_q(X, D)B_p(X, D) = B_q(X, D)A_p(X, D) = |X|.$$

C. Outer distribution and fundamental parameters of codes

The inner distribution of a code Y is concerned with the mutual relations or "distances" between the code points which are values of the function $\partial_R(x, y)$. We shall omit the quotes when, for a symmetric association scheme (X, R), $\partial_R(x, y)$ satisfies the triangle inequality and hence is a distance function. Let $D(Y)$ denote the set of distinct values of the function $\partial_R(x, y)$ when $x, y \in Y$, $x \neq y$. Note that

$$D(Y) = \{i \in N_n^1 \;:\; a_i(Y) \neq 0\}$$

and define

$$D'(Y) := \{i \in N_n^1 \;:\; a'_i(Y) \neq 0\}.$$

For simplicity we shall consider codes Y in an n-class association scheme (X, R), such that $1 < |Y| < |X|$. Then we can state that both $D(Y)$ and $D'(Y)$ are not empty. Define the following fundamental parameters of a code Y [36]:

- *the minimum "distance"* $d(Y) := \min D(Y)$
- *the (minimum) dual "distance"* $d'(Y) := \min D'(Y)$
- *the degree* $s(Y) := |D(Y)|$
- *the dual degree* $s'(Y) := |D'(Y)|$

Together with $d'(Y)$ we will also consider

- *the (maximum) strength* $\tau(Y) := d'(Y) - 1$.

Moreover, we consider two auxiliary parameters:

$$\gamma(Y) = \begin{cases} 1 \text{ if } n \in D(Y) \\ 0 \text{ otherwise,} \end{cases} \quad \text{and } \gamma'(Y) = \begin{cases} 1 \text{ if } n \in D'(Y) \\ 0 \text{ otherwise.} \end{cases}$$

For given integers d and τ (with $1 \leq d \leq n$ and $0 \leq \tau \leq n-1$), a code Y is called a *d-code* if $d(Y) \geq d$ and a *τ-design* if $\tau(Y) \geq \tau$. These notions are special cases of a D-code and a D-design, respectively, for $D = N_n^d$ and $D = N_n^{\tau+1}$. The examples of τ-designs in the Hamming and Johnson schemes are examined in Section III-A above. The given definitions clearly show the dual character of the notions of d-codes and $(d-1)$-designs. Let $A(X, d)$ denote the maximum size of a d-code in (X, R) and let $B(X, d)$ denote the minimum size of a $(d-1)$-design in (X, R). A d-code Y in (X, R) is called *maximal* if $|Y| = A(X, d)$ and a $(d-1)$-design Y in (X, R) is called *minimal* if $|Y| = B(X, d)$.

Now we introduce a definition that involves the relations ("distances") between the code Y and the whole ambient set X.

Definition 8: The *outer distribution* of a code Y in an n-class association scheme (X, R) is the $|X| \times (n+1)$ matrix M whose (x, i) entry $M_i(x)$ equals

$$a_i(x, Y) := |\{y \in Y : (x, y) \in R_i\}|.$$

Some fundamental properties of a code Y are defined in terms of the rows $M(x) = (a_0(x, Y), ..., a_n(x, Y))$ of its outer distribution M. A code Y is called *distance invariant* if $M(x) = M(y)$ for any $x, y \in Y$ and *completely distance invariant* if $M(x) = M(y)$ for any $x, y \in X$ such that $\partial_R(x, Y) = \partial_R(y, Y)$ where $\partial_R(x, Y) = \min\{\partial_R(x, z) : z \in Y\}$. When Y is a *linear code over* $\mathbb{F}_q$ (in the Hamming scheme), the rows $M(x)$ of the outer distribution M are the weight distributions of the coset codes $Y + x$.

It is easily seen that $M^T M$ can be expressed linearly in terms of the inner distribution of Y. Let us give the "Q-transform version" of this expression. Consider the Q-transform of each row $M(x)$ of the outer distribution M. This produces the matrix MQ, with $Q := (q_k(i))_{i,k \in N_n}$.

Theorem 7: [37] The Q-transform MQ of the outer distribution is related to the Q-transform $(a'_k(Y))_{k=0}^n$ of the inner distribution by

$$Q^* M^T M Q = |X||Y| \operatorname{diag}(a'_0(Y), a'_1(Y), \cdots, a'_n(Y)).$$

As an immediate consequence, the *rank of M is equal to* to $s'(Y) + 1$. Furthermore, we obtain $\sum_{x \in X} |a'_k(x, Y)|^2 = |X||Y| a'_k(Y)$, whence $a'_k(x, Y) = 0$ for any $k \in N_n^1 \setminus D'(Y)$.

This can also be deduced from (34) as follows: If $k \in N_n^1 \setminus D'(Y)$, then

$$a'_k(x,Y) = \sum_{i=0}^{n} a_i(x,Y) q_k(i) = \sum_{y\in Y} q_k(\partial_R(x,y)) = \sum_{j=1}^{m_k} v_{k,j}(x) \sum_{y\in Y} v_{k,j}(y) = 0. \quad (40)$$

Some interesting problems in classical coding theory are concerned with the *covering radius* $\rho(Y)$ of a code Y (in the Hamming scheme) (see [6], [31], [34], [64]). By definition, $\rho(Y) := \max\{\partial_H(x,Y) : x \in X\}$ where $\partial_H(x,Y) := \min\{\partial_H(x,y) : y \in Y\}$. (Thus, for a linear code, $\rho(Y)$ is the maximum weight of coset leaders.) This definition is extended to any association scheme (X,R) if one replaces $\partial_H(x,u)$ by $\partial_R(x,u)$. The covering radius $\rho(Y)$ can be found from the outer distribution M of Y, since $\partial_R(x,Y)$ is the smallest $i \in N_n$ such that $a_i(x,Y) \neq 0$. Notice that $\rho(Y)$ generally cannot be determined from the inner (distance) distribution of Y; however, some upper bounds on $\rho(Y)$ can be obtained from these data [37], [51], [118], [127]. (See also Sections IV-C and IV-D below.)

IV. Polynomial schemes

A. Orthogonal polynomials

In the examples of the Hamming and Johnson schemes (and in several other interesting cases), the p-numbers $p_i(k)$ and the q-numbers $q_k(i)$ are representable by polynomials of degree i and k, respectively, in an "appropriate variable" (see Section II-B). This leads us to investigate the class of association schemes that enjoy either of these "polynomial properties" (or both of them).

Definition 9: (i) A symmetric n-class association scheme is *P-polynomial*, with respect to a function $\sigma_P \in F(N_n)$, if there exist real polynomials $P_i(t)$ of degree $i, i = 0,1,...,n$, such that $p_i(k) = P_i(\sigma_P(k))$ for any $k \in N_n$.

(ii) A symmetric n-class association scheme is *Q-polynomial*, with respect to a function $\sigma_Q \in F(N_n)$, if there exist real polynomials $Q_k(t)$ of degree $k, k = 0,1,...,n$, such that $q_k(i) = Q_k(\sigma_Q(i))$ for any $i \in N_n$.

It can be proved that these functions σ_P and σ_Q must be linear functions of the first p- and q-functions p_1 and q_1 (respectively), which take different values $p_1(j)$ and $q_1(j)$ for different $j \in N_n$ such that $|p_1(j)| \le p_1(0) = v_1$ and $|q_1(j)| \le q_1(0) = m_1$. We will use the following functions σ_P and σ_Q:

$$\sigma_P(d) = 1-2\frac{p_1(d)-v_1}{p_1(n)-v_1}, \quad \sigma_Q(d) = 1-2\frac{q_1(d)-m_1}{q_1(n)-m_1}. \quad (41)$$

Then $\sigma_P(0) = 1, \sigma_P(n) = -1$ and $\sigma_Q(0) = 1, \sigma_Q(n) = -1$. When the function p_1 or q_1 is decreasing on N_n we will extend it to a continuous decreasing function on $[0,n]$ (usually the latter is defined by the same formula). In this case the corresponding function given by (41) is a decreasing continuous mapping from $[0,n]$ onto $[-1,1]$ and it is called *standard*.

It follows directly from Theorem 1 and (15) that $P = \{P_i(t) : i \in N_n\}$ and $Q = \{Q_k(t) : k \in N_n\}$ are *systems of orthogonal polynomials* with the following orthogonality conditions:

$$\sum_{k=0}^{n} P_i(\sigma_P(k)) P_j(\sigma_P(k)) m_k = v_i \, |X| \, \delta_{i,j}, \quad (42)$$

$$\sum_{i=0}^{n} Q_k(\sigma_Q(i)) Q_l(\sigma_Q(i)) v_i = m_k \, |X| \, \delta_{k,l}. \quad (43)$$

and the properties:

$$m_k P_i(\sigma_P(k)) = v_i Q_k(\sigma_Q(i)), \quad i,k \in N_n. \quad (44)$$

These orthogonal systems P and Q are uniquely determined by *three-term recurrence relations* [49] of the form

$$p_{i,1}^{i+1} P_{i+1}(t) = \left(P_1(t) - p_{i,1}^{i}\right) P_i(t) - p_{i,1}^{i-1} P_{i-1}(t),$$
$$q_{k,1}^{k+1} Q_{k+1}(t) = \left(Q_1(t) - q_{k,1}^{k}\right) Q_k(t) - q_{k,1}^{k-1} Q_{k-1}(t),$$

where

$$P_0(t) = 1, \qquad 2P_1(t) = v_1 + p_1(n) + t(v_1 - p_1(n)),$$
$$Q_0(t) = 1, \qquad 2Q_1(t) = m_1 + q_1(n) + t(m_1 - q_1(n)).$$

Definition 9 depends on the ordering of the relations R_i and of the idempotents E_k, respectively. For this reason, a given association scheme may possess more than one P-polynomial structure and more than one Q-polynomial structure (see [10]).

The algebraic notion of a P-polynomial scheme is equivalent [37] to the combinatorial notion of a distance-regular graph, defined as follows. Let (X, R_1) be a simple connected finite graph of diameter n. For $i \in N_n$, define R_i as the set of pairs $(x,y) \in X^2$ such that x and y are at distance i apart in (X,R_1), and let $R := \{R_0, R_1, \cdots, R_n\}$. If (X,R) is an association scheme, then (X,R_1) is said to be *distance-regular* [26]. Thus, if (X,R) is a P-polynomial scheme, then ∂_R (see (1)) is a distance function. Note that a symmetric association scheme (X,R) with a distance function ∂_R need not be P-polynomial. (An example is provided by the "ordered Hamming scheme" [89].) On the other hand, the algebraic notion of a Q-polynomial scheme has no simple combinatorial interpretation, except in some important special cases where (X,R) can be embedded in a certain "lattice-type structure" [38], [39], [94]. Nevertheless, there exist some useful general characterizations of Q-polynomial schemes [54], [126]. There is an elementary criterion for the P-polynomial property in terms of the intersection numbers, namely $p_{i,1}^{i+1} \neq 0$ and $p_{i,1}^{k} = 0$ for $k > i+1$. This characterizes the "distance structure" in a clear manner. Similarly, a criterion for the Q-polynomial property is $q_{k,1}^{k+1} \neq 0$ and $q_{k,1}^{t} = 0$ for $t > k+1$.

Example 1 (continued) For the Hamming scheme H_q^n, (17) holds and $p_1(d) = q_1(d) = n(q-1) - dq$. Hence, H_q^n is P- and Q-polynomial with respect to the standard function $\sigma(d) = 1 - 2d/n$, systems P and Q coincide and consist of the polynomials $P_i(t) = K_i^n((1-t)n/2), i = 0,1,...,n$.

Example 2 (continued) For the Johnson scheme J_n^v, (20) holds and $p_1(d) = v_1 - d(v+1-d)$, $q_1(d) = m_1(1 - \frac{dv}{n(v-n)})$. Hence, J_n^v is P-polynomial with respect to the standard function $\sigma_P(d) = 1 - 2\frac{d(v+1-d)}{n(v+1-n)}$ and Q-polynomial with respect to the standard function $\sigma_Q(d) = 1 - 2d/n$. Systems P and Q are defined by means of $P_i(\sigma_P(z)) = \tilde{H}_i(z)$ and $Q_i(\sigma_Q(z)) = H_i(z)$, $i = 0, 1, ..., n$.

The number of independent parameters of an n-class P-polynomial or Q-polynomial scheme is equal to $2n-1$. In contrast with this observation, Leonard has proved that all parameters of a P- *and* Q-polynomial scheme can be determined from only *five* independent numbers [72]. The same author [71] has shown that the polynomials $P_i(t)$ and $Q_k(t)$ relative to P- and Q-polynomial schemes belong to a well-defined 5-parameter class of orthogonal polynomials of the generalized hypergeometric type [114], known as the *Askey–Wilson polynomials* [2]. His result produces closed form expressions for the p-numbers and the q-numbers. Furthermore, it characterizes the Askey–Wilson polynomials as those orthogonal polynomials having "duals".

For a code Y in a Q-polynomial scheme, a polynomial f is called an *annihilator* for Y if $f(\sigma_Q(i)) = 0$ for all $i \in D(Y)$. An annihilator f of minimal degree (i.e., degree $s(Y)$) is called *minimal* and denoted by $f^{Y,Q}$ if $f(\sigma_Q(0)) = f(1) = 1$. For a code Y in a P-polynomial scheme, a polynomial f is called a *dual annihilator* for Y if $f(\sigma_P(i)) = 0$ for all $i \in D'(Y)$. A dual annihilator f of minimal degree (i.e., degree $s'(Y)$) is called *dual minimal* and denoted by $f^{Y,P}$ if $f(\sigma_P(0)) = f(1) = 1$. In particular, if for $d \in N_n$ and $U = P$ or $U = Q$,

$$g^{d,U}(t) = \prod_{j=d}^{n} \frac{t - \sigma_U(j)}{1 - \sigma_U(j)}, \tag{45}$$

then $f^{X,Q} = g^{0,Q}$ and $f^{X,P} = g^{0,P}$. For any non-empty $D \subseteq N_n$ and any $i \in D$ denote by $g^{D,i,U}$ the polynomial of degree $|D|-1$ uniquely defined by the conditions: $g^{D,i,U}(\sigma_U(j)) = \delta_{i,j}$ for any $j \in D$. Any function h on D can be represented by the *interpolation* polynomial $\sum_{i=0}^n h(i) g^{D,i,U}$ of degree $|D|-1$. In particular, for any $h \in F(N_n)$ we have $h(j) = \sum_{i=0}^n h(i) g^{N_n,i,U}(\sigma_U(j))$.

In the case of a U-polynomial association scheme (X, R) (where U is either P or Q), we can rephrase the linear programming bounds of Theorems 4 and 5 in terms of extremum problems for the system U of orthogonal polynomials. Denote by $F_n[t]$ the set of real polynomials of degree at most n in t. For any $f \in F_n[t]$, let $f_j(U)$, $j \in N_n$, be the coefficients of the (unique) expansion of f over the system U, i.e., $f = \sum_{j=0}^n f_j(U) U_j$. Put $\Omega_U(f) = f(1)/f_0(U)$ if $f_0(U) \neq 0$. Note that $f \to h := f(\sigma_U)$ gives a one-to-one mapping of $F_n[t]$ onto $F(N_n)$ with $h(j) = f(\sigma_U(j))$ and $h_j(u) = f_j(U)$ for any $j \in N_n$ (see (14) and Definition 9). We restrict our attention to the case of d-codes and $(d-1)$-designs (i.e., codes with dual "distance" d or more), which corresponds to the case of D-codes and D-designs for $D = N_n^d$. We say that $f \in F_n[t]$ has the *property* $\mathfrak{A}_U(d)$ ($\mathfrak{B}_U(d)$) if $h = f(\sigma_U) \in F(N_n)$ has the property $\mathfrak{A}_u(N_n^d)$ (respectively $\mathfrak{B}_u(N_n^d)$). Let $A_U(X, d) = \min \Omega_U(f)$ where the minimum is taken over all polynomials $f \in F_n[t]$ with the property $\mathfrak{A}_U(d)$. Similarly, let $B_U(X, d) = \max \Omega_U(f)$ where the maximum is taken over all polynomials $f \in F_n[t]$ with the property $\mathfrak{B}_U(d)$. Since $\Omega_U(f) = \Omega_u(h)$ for $h = f(\sigma_U)$, we have

$$A_U(X, d) = A_u(X, N_n^d), \quad B_U(X, d) = B_u(X, N_n^d) \tag{46}$$

and we can use Theorems 4, 5 and 6 to estimate the size of d-codes and $(d-1)$-designs with the help of the above extremum problems for the system U.

Without going into any detail, let us finally point out that the classical examples of P-and Q-polynomial schemes are induced by some classical *permutation groups* (see Section II-D). Extensive research work has been devoted, in this context, to the subject of "orthogonal polynomials and permutation groups" [48], [119].

B. Adjacent systems of orthogonal polynomials and two extremum problems

Some important estimates on fundamental parameters of codes are expressed in terms of values connected with systems of orthogonal polynomials which are *adjacent* to the systems P and Q. Consider two functions σ and w on N_n. We assume that the first function (*change of variable*) σ takes the values $\sigma(n) = -1$, $\sigma(0) = 1$, and maps N_n into the interval $[-1, 1]$, and the second (*weight*) function w has the properties $w(i) > 0$ and $\sum_{i=0}^n w(i) = 1$. We call the change of variable σ *standard* if it can be represented as a continuous decreasing function on the whole interval $[0, n]$. It is known [49], [121] that the orthogonality conditions

$$\sum_{d=0}^{n} U_i(\sigma(d))\, U_j(\sigma(d))\, w(d) = U_i(1)\, \delta_{i,j} \tag{47}$$

uniquely define a system $U = \{U_i(t) : i \in N_n\}$ of polynomials $U_i(t)$ of degree i with some positive values $U_i(1)$. We denote by σ_U and w_U the functions σ and w for the system U. In particular, for the systems P and Q we have (see (42) and (43)) $w_P(i) = |X|^{-1} m_i$, $P_i(1) = v_i$, $w_Q(i) = |X|^{-1} v_i$, $Q_i(1) = m_i$. We assume that U satisfies the *Krein condition*: for any $i, j, k \in N_n$ there exist nonnegative real numbers $u_{i,j}^k$ such that

$$U_i(t) U_j(t) = \sum_{k=0}^{n} u_{i,j}^k U_k(t) \pmod{g^{0,U}(t)}.$$

By (23) and (24) this is fulfilled for the systems P and Q. For the system U and any $k \in N_n$ we define the *kernel* function

$$T_k(t_1, t_2; U) = \sum_{i=0}^{k} U_i(t_1) U_i(t_2) / U_i(1).$$

For any $a, b \in \{0, 1\}$ we consider a *new* weight function $w_U^{a,b}$ on N_n such that

$$w_U^{a,b}(i) = c^{a,b} (1 - \sigma_U(i))^a (1 + \sigma_U(i))^b\, w_U(i) \tag{48}$$

where the constant $c^{a,b}$ is chosen so that $\sum_{i=0}^{n} w_U^{a,b}(i) = 1$. The initial change of variable σ_U and the new weight function $w_U^{a,b}$ uniquely define a system $U^{a,b} = \{U_i^{a,b}(t) : i \in N_{n-a-b}\}$ of polynomials $U_i^{a,b}(t)$ of degree i by means of the following conditions:

$$\sum_{d=0}^{n-a-b} U_i^{a,b}(\sigma_U(d))\, U_j^{a,b}(\sigma_U(d))\, w_U^{a,b}(d) = U_i^{a,b}(1)\delta_{i,j}. \tag{49}$$

(The system $U^{a,b}$ consists of $n+1-a-b$ polynomials since $a+b$ weights become zero.) We put $T_k^{a,b}(t_1,t_2;U) := T_k(t_1,t_2;U^{a,b})$. Let $t_k^{a,b}(U)$ be the largest zero of the polynomial $U_k^{a,b}$. If σ_U is standard we can uniquely define the numbers $d_k^{a,b}(U)$ by $\sigma_U\left(d_k^{a,b}(U)\right) = t_k^{a,b}(U)$. We will omit the indices a,b in the notations $U_k^{a,b}$, $t_k^{a,b}(U)$, $d_k^{a,b}(U)$ when $a=b=0$.

Example 1 (continued) Let $K_k^n(z)$ be the Krawtchouk polynomial of degree k defined by (16) and let $d_k(n)$ be its smallest zero. For the Hamming scheme H_q^n,

$$P_k^{a,b}(1-2z/n) = C_k^{a,b} K_k^{n-a-b}(z-a)$$

where $C_k^{1,b} = \sum_{i=0}^{k}\binom{n-b}{i}(q-1)^{i-k}/\binom{n-1-b}{k}$ and $C_k^{0,b} = 1$, and hence

$$d_k^{a,b}(Q) = d_k^{a,b}(P) = d_k(n-a-b)+a. \tag{50}$$

In particular, $Q_k^{0,1}(1) = \binom{n-1}{k}(q-1)^k$.

Example 2 (continued) Let $H_k^{v,n}(z)$ and $\tilde{H}_k^{v,n}(z)$ be the polynomials of degree k defined respectively by (18) and (19), and let $d_k(v,n)$ and $\tilde{d}_k(v,n)$ be their smallest zeros. For the Johnson scheme J_n^v,

$$Q_k^{0,1}(\sigma_Q(z)) = H_k^{v-1,n-1}(z)$$

and $P_k^{1,0}(\sigma_P(z))$ is proportional to $\tilde{H}_k^{v-2,n-1}(z-1)$. Hence $d_k^{0,1}(Q) = d_k(v-1,n-1)$, $d_k^{1,0}(P) = \tilde{d}_k(v-2,n-1)+1$, and $Q_k^{0,1}(1) = \binom{v-1}{k} - \binom{v-1}{k-1}$.

Now we consider two extremum problems for the system U of orthogonal polynomials under consideration. For any $d \in N_n^2$, the $K_U(d)$*-problem* consists in finding

$$K_U(d) := \max \Omega_U(f)$$

where the maximum is taken over all polynomials $f \in F_{d-1}[t]$ such that $f_0(U) > 0$ and $f(t) \ge 0$ for $-1 \le t \le 1$. A polynomial f having these properties for which $\Omega_U(f) = f(1)/f_0(U) = K_U(d)$ is called an *optimal solution* of the $K_U(d)$-problem. Note that these properties are in general stronger than $\mathfrak{B}_U(d)$ since they include nonnegativity f on the whole interval $[-1,1]$ (not only at points $\sigma_U(i)$, $i \in N_n$) and a restriction on its degree. This implies that for any U-polynomial (U is either P or Q) association scheme X,

$$B_U(X,d) \ge K_U(d). \tag{51}$$

From now on we assume that l and θ denote arbitrary numbers such that $l \in N_n^1$ and $\theta \in \{0,1\}$.

Theorem 8: [103] For any $d = 2l+1-\theta \in N_n^2$ the polynomial

$$g^{(d)}(t) = (t+1)^\theta \left(U_{l-\theta}^{1,\theta}(t)\right)^2$$

of degree $d-1$ is the unique (up to a constant factor) optimal solution of the $K_U(d)$-problem and

$$K_U(d) = \left(1 - \frac{U_1(1)}{U_1(-1)}\right)^\theta \sum_{i=0}^{l-\theta} U_i^{0,\theta}(1).$$

One can show that $K_U(d)$ is a positive-valued increasing function in $d \in N_n$ and admits another expression:

$$K_U(2l+1-\theta) = \left(1 - \frac{U_l(1)U_{l-1}^{1,0}(-1)}{U_l(-1)U_{l-1}^{1,0}(1)}\right)^\theta \sum_{i=0}^{l-\theta} U_i(1).$$

For odd d and $U = Q$ the polynomials $g^{(2l+1)}(t) = \left(\sum_{i=0}^{l} U_i(t)\right)^2$ were first used in 1973 in [37] to obtain a lower bound on the size of $(d-1)$-designs (see Theorem 19 below). In the general case Theorem 8 was applied to this end in [47].

Example 1 (continued) For the Hamming scheme H_q^n and $d = 2l+1-\theta$,

$$K_P(d) = K_Q(d) = q^\theta \sum_{i=0}^{l-\theta}\binom{n-\theta}{i}(q-1)^i. \tag{52}$$

Example 2 (continued) For the Johnson scheme J_n^v and $d = 2l+1-\theta$,

$$K_P(d) = \sum_{i=0}^{l-\theta}\binom{n-\theta}{i}\binom{v-n+\theta}{i+\theta}, \tag{53}$$

$$K_Q(d) = \left(\frac{v}{n}\right)^\theta \binom{v-\theta}{l-\theta}. \tag{54}$$

Now we formulate the second extremum problem for the system U with a standard function σ_U. It is known [103] that the largest zeros $t_k^{a,b} = t_k^{a,b}(U)$ of the polynomials $U_k^{a,b}$ satisfy the following inequalities:

$$t_{k-1}^{1,1} < t_k^{1,0} < t_k^{1,1}, \quad k = 1,\dots,n-1,$$

where it is assumed that $t_0^{1,1} = -1 = \sigma_U(n)$ and $t_{n-1}^{1,1} = \sigma_U(1)$. This means that the half-open interval $[-1,\sigma_U(1))$ is partitioned into the half-open intervals $[t_{k-1}^{1,1}, t_k^{1,0})$ and $[t_k^{1,0}, t_k^{1,1})$, $k = 1,\dots,n-1$. Enumerate in succession all these half-open intervals from the left to the right by positive integers. For any real number σ, $-1 \le \sigma < \sigma_U(1)$, denote by $h(\sigma)$ *the number* of the (unique) half-open interval containing σ. Let $k(\sigma) = k$ when $\sigma \in [t_{k-1}^{1,1}, t_k^{1,0})$ or $\sigma \in [t_k^{1,0}, t_k^{1,1})$, and let $\varepsilon(\sigma) = 0$ if $\sigma \in [t_{k(\sigma)-1}^{1,1}, t_{k(\sigma)}^{1,0})$ and $\varepsilon(\sigma) = 1$ if $\sigma \in [t_{k(\sigma)}^{1,0}, t_{k(\sigma)}^{1,1})$. Then it is clear that $h(\sigma) = 2k(\sigma) - 1 + \varepsilon(\sigma)$. For any number σ, $-1 \le \sigma < \sigma_U(1)$, the $L_U(\sigma)$*-problem* consists in finding

$$L_U(\sigma) := \min \Omega_U(f)$$

where the minimum is taken over all polynomials $f \in F_{h(\sigma)}[t]$ such that $f_0(U) > 0$ and $f(t) \leq 0$ for $-1 \leq t \leq \sigma$. A polynomial f having these properties for which $\Omega_U(f) = f(1)/f_0(U) = L_U(\sigma)$ is called an *optimal solution* of the $L_U(\sigma)$-problem. Note that for $\sigma = \sigma_U(d)$ these properties as compared to $\mathfrak{A}_U(d)$ say nothing about nonnegativity of $f_i(U)$ for $i \in N_n^1$, include a stronger condition than $f(\sigma_U(i)) \leq 0$ for $i \in N_n^d$ and introduce a restriction to the degree of the polynomials. Note that this restriction means that, in the $L_U(\sigma)$-problem, polynomials whose degree does not exceed the number of the half-open interval containing σ are considered. This also holds in the case of the $K_U(d)$-problem since $(1, n]$ is partitioned into the half-open intervals $(i, i+1]$, $i \in N_{n-1}^1$, and $d-1$ is the number of the half-open interval containing $d \in N_n^2$.

Theorem 9: [77], [81] For any real number σ, $-1 \leq \sigma < \sigma_U(1)$, let $\varepsilon = \varepsilon(\sigma)$ and $k = k(\sigma)$. Then the polynomial

$$f^{(\sigma)}(t) = (t-\sigma)(t+1)^{\varepsilon}\left(T_{k-1}^{1,\varepsilon}(t,\sigma;U)\right)^2 \tag{55}$$

of degree $h(\sigma) = 2k-1+\varepsilon$ is an optimal solution of the $L_U(\sigma)$-problem. The function $L_U(\sigma)$ is equal to

$$\left(1 - \frac{U_1(1)}{U_1(-1)}\right)^{\varepsilon}\left(1 - \frac{U_k^{0,\varepsilon}(1)U_{k-1}^{1,\varepsilon}(\sigma)}{U_k^{0,\varepsilon}(\sigma)U_{k-1}^{1,\varepsilon}(1)}\right)\sum_{i=0}^{k-1} U_i^{0,\varepsilon}(1),$$

positive-valued and continuous, grows with σ and takes the following values at the left ends of these half-open intervals

$$L_U(t_{l-\theta}^{1,\theta}) = K_U(2l+1-\theta). \tag{56}$$

We give some additional facts on the polynomials $f^{(\sigma)}(t)$. For $\sigma \neq t_{l-\theta}^{1,\theta}$ the polynomial $f^{(\sigma)}(t)$ is the unique (up to a constant factor) optimal solution of the $L_U(\sigma)$-problem. For $\sigma = t_l^{1,0}$ we have $k(\sigma) = l$, $\varepsilon(\sigma) = 1$ and the polynomial $f^{(\sigma)}(t)$ has factor $t+1$. For $\sigma = t_{l-1}^{1,1}$ we have $k(\sigma) = l$, $\varepsilon(\sigma) = 0$, and $f^{(\sigma)}(t)$ is also divisible by $t+1$. In both cases the polynomial $f(t) = f^{(\sigma)}(t)/(t+1)$ is an optimal polynomial for the $L_U(\sigma)$-problem as well. Moreover, for $\sigma = t_{l-\theta}^{1,\theta}$ the polynomial $(t-\sigma)\, f^{(\sigma)}(t)/(t+1)$ is proportional to the optimal solution $g^{(2l+1-\theta)}(t)$ of the $K_U(2l+1-\theta)$-problem. These facts and Theorems 8 and 9 follow from the following main theorem which (as we shall see below) also determines the inner distribution of optimal codes and designs.

Theorem 10: [77], [81] For any σ, $-1 \leq \sigma < \sigma(1)$, the polynomial

$$(t-\sigma)(t+1)^{\varepsilon}T_{k-1}^{1,\varepsilon}(t,\sigma;U) \tag{57}$$

with $k = k(\sigma)$ and $\varepsilon = \varepsilon(\sigma)$ has $k+\varepsilon$ simple zeros $\alpha_0, \alpha_1, \ldots, \alpha_{k+\varepsilon-1}$ ($\alpha_0 < \alpha_1 < \ldots < \alpha_{k+\varepsilon-1}$) where $\alpha_{k+\varepsilon-1} = \sigma$ and $\alpha_0 \geq -1$ with equality holding if and only if $\varepsilon = 1$ or $\varepsilon = 0$ and $\sigma = t_{k-1}^{1,1}$. Moreover, for any polynomial $f(t)$ of degree at most $h(\sigma) = 2k-1+\varepsilon$ the following equality holds:

$$f_0(U) = (L_U(\sigma))^{-1}f(1) + \sum_{j=0}^{k+\varepsilon-1}\rho_j^{(\sigma)}(U)f(\alpha_j) \tag{58}$$

where for $i = 0, ..., k-1$,

$$\rho_{i+\varepsilon}^{(\sigma)}(U) = \frac{1}{c^{1,\varepsilon}\,(1+\alpha_{i+\varepsilon})^{\varepsilon}\,(1-\alpha_{i+\varepsilon})\,T_{k-1}^{1,\varepsilon}(\alpha_{i+\varepsilon},\alpha_{i+\varepsilon};U)}$$

are positive, and in the case $\varepsilon = 1$,

$$\rho_0^{(\sigma)}(U) = \frac{T_k(\sigma,1)}{T_k(-1,-1)T_k(\sigma,1) - T_k(-1,1)T_k(\sigma,-1)} \geq 0$$

with equality holding if and only if $\sigma = t_k^{1,0}$ (here U is omitted in the notation $T_k(t_1, t_2; U)$).

Example 1 (continued) In the case of the Hamming scheme H_q^n for any $d \in N_n^2$ there exist $k \in N_{n-1}^1$ and $\varepsilon \in \{0,1\}$ such that $d_k(n-1-\varepsilon) < d-1 \leq d_{k-1+\varepsilon}(n-2+\varepsilon)$. Then $L_Q(\sigma) = L_P(\sigma)$ for $\sigma = \sigma(d) = 1 - 2d/n$ can be expressed by the following formula:

$$q^{\varepsilon}\left(\sum_{i=0}^{k-1}\binom{n'}{i}(q-1)^i - \binom{n'}{k}(q-1)^k\frac{K_{k-1}^{n'-1}(d-1)}{K_k^{n'}(d)}\right),$$

where $n' = n - \varepsilon$. In particular, when the number d belongs to the half-open intervals

$$\left(\tfrac{(q-1)n+1}{q}, n\right], \quad \left(\tfrac{(q-1)(n-1)+1}{q}, \tfrac{(q-1)n+1}{q}\right],$$

$$\left(d_2(n-1)+1, \tfrac{(q-1)(n-1)+1}{q}\right]$$

this respectively gives the values

$$\tfrac{qd}{qd-(q-1)n}, \quad \tfrac{q^2 d}{qd-(q-1)(n-1)},$$

$$\tfrac{qd((n(q-1)+1)(n(q-1)-qd+2)-q)}{qd(2n(q-1)-q+2-qd)-n(n-1)(q-1)^2}$$

which are obtained with the help of the optimal polynomials $f^{(\sigma)}(t)$ of the first, second and third degree.

Example 2 (continued) In the case of the Johnson scheme J_n^v, when d belongs to the half-open intervals

$$\left(\tfrac{n(v-n)}{v-1}, n\right], \left(\tfrac{(n-1)(v-n)}{v-2}, \tfrac{n(v-n)}{v-1}\right], \left(d_2^{1,0}(Q)+1, \tfrac{(n-1)(v-n)}{v-2}\right]$$

the function $L_Q(\sigma)$ for $\sigma = \sigma_Q(d) = 1 - 2d/n$ respectively equals

$$\tfrac{vd}{vd-n(v-n)}, \quad \tfrac{v(v-1)d}{n(d(v-1)-(n-1)(v-n))},$$

$$\tfrac{v(v-1)d(n(v-n)-1-d(v-2))}{d(v-1)(2n(v-n)-v-d(v-2))-n(n-1)(v-n)(v-n-1)}.$$

This is obtained with the help of the optimal polynomial $f^{(\sigma)}(t)$ for the $L_Q(\sigma)$-problem of the first, second and third degree (see [76]).

In order to prove that for $\sigma = \sigma_U(d)$ the optimal polynomial $f^{(\sigma)}(t)$ for the $L_U(\sigma)$-problem has the property $\mathfrak{A}_U(d)$ and hence the inequality $A_U(X, d) \leq L_U(\sigma_U(d))$ holds, one must check that all coefficients $f_i^{(\sigma)}(U)$ of its expansion over system U are non-negative. Note that in the case $U = Q$ by Theorem 2 the latter means that the symmetric matrix $f^{(\sigma)}(\sigma_Q(\partial_R(x,y)))$ for Q-polynomial association scheme (X, R) is nonnegative definite. Now it is known

[77], [81] that all coefficients $f_i^{(\sigma)}(U)$, $i \in N_{h(\sigma)}$, are positive when $\varepsilon(\sigma) = 0$ (or $f^{(\sigma)}$has odd degree $2k(\sigma) - 1$), in particular, for $\sigma = t_{l-1}^{1,1}$, and also when $\sigma = t_l^{1,0}$, $l \in N_n^1$. Moreover, the same is true for all σ if the system U satisfies the *strengthened* Krein condition: for any $i, j \in N_{n-2}$ the coefficients of the expansion of $(1 + t)U_i^{1,1}(t)U_j^{1,1}(t)$ $(\bmod\, g^{0,U}(t))$ over the system U are positive. It is known that the system Q satisfies the strengthened Krein condition for *decomposable* Q-polynomial schemes [77]. The class of decomposable schemes contains some known infinite families of P- and Q-polynomial association schemes, in particular, H_q^n and J_n^v. It seems to be true that all coefficients $f_i^{(\sigma)}(U)$, $i \in N_{h(\sigma)}$, are also positive when σ belongs to the open interval $(t_{k(\sigma)}^{1,0}, t_{k(\sigma)}^{1,1})$. Unfortunately, this question is still open for $k(\sigma) \geq 2$. Thus, for any system U under consideration,

$$A_U(X, d) \leq L_U(t_{l-\theta}^{1,\theta}) = K_U(2l + 1 - \theta) \tag{59}$$

if $d \geq d_{l-\theta}^{1,\theta}(U)$ (see (56)), and for any d,

$$A_U(X, d) \leq L_U(\sigma_U(d)). \tag{60}$$

if U satisfies the strengthened Krein condition.

The known earlier bounds for a d-code Y can be described in terms of polynomials $f(t)$ which have the $\mathfrak{A}_Q(d)$-property ($f_0(Q) > 0$, $f_i(Q) \geq 0$ for $i = 1, ..., n$, $f(1) > 0$, $f(t) \leq 0$ for $-1 \leq t \leq \sigma = \sigma_Q(d)$) and hence imply $|Y| \leq \Omega(f)$. In particular, the bounds due to Blichfeldt [19], Rankin [97], Plotkin and Johnson (see [88]) are in fact based on the polynomial $f(t) = t - \sigma$ with $\sigma = \sigma_Q(d) \leq t_1(Q)$ which provides $f_0(Q) > 0$, $f_1(Q) \geq 0$. The Sidelnikov results [110], [111] are obtained with the help of polynomials $f(t) = t^{2l} - \sigma^{2l}$ for a suitable choice of an integer l. In [90] (and later in [65] for the Euclidean sphere) the polynomials

$$f(t) = (t - \sigma)\left(T_{k-1}(t, \sigma; Q)\right)^2 \tag{61}$$

were used, where the integer k is defined by $t_{k-1} < \sigma = \sigma_Q(d) < t_k$. The polynomials (55) were found with the help of the Lagrange method and presented in [73]. Some extremum properties of these polynomials were found in [112] and they were essentially used to prove the optimality of (55) for the $L_U(\sigma)$-problem (see Theorems 10 and 9).

The solutions of the $K_U(d)$- and $L_U(\sigma)$-problems can also be applied to codes and designs in the Cartesian product X^m of m copies of a P-polynomial association scheme (X, R) with the distance $\max_{1 \leq i \leq m} \partial_R(x_i, y_i)$ or $\sum_{i=1}^m \partial_R(x_i, y_i)$ on X^m. In particular, for the case of the distance $\max_{1 \leq i \leq m} \partial_R(x_i, y_i)$ on X^m, this allows one to estimate the *Shannon capacity* [106] $\lim_{m \to \infty} \left(A(X^m, 2)\right)^{1/m}$ of a graph (X, R_1) (see [85], [91], [104], and also [79]).

In the remainder of this section, considering a code Y in an n-class P- and/or Q-polynomial scheme we shall assume for simplicity that $n \geq 2$, $d(Y) \geq 2$, $d'(Y) \geq 2$.

C. Codes and designs in P-polynomial schemes

Throughout this subsection, we consider an n-class P-polynomial scheme (X, R). In this case $\partial_R(x, y)$ (see (1)) is a distance function and (X, R) can be considered as a metric space X with the metric $\partial(x, y) = \partial_R(x, y)$. It follows that for any code $Y \subseteq X$, the *metric spheres* (balls) $S_r(y) := \{x \in X : 0 \leq \partial(x, y) \leq r\}$ of radius r centered at the code points $y \in Y$ do not intersect when r is equal to the *packing radius* $e(Y) = \lfloor (d(Y) - 1)/2 \rfloor$ and cover X when r is equal to the covering radius $\rho(Y)$. This gives the following *sphere packing and sphere covering bounds* for any code Y in (X, R):

$$|Y| \sum_{i=0}^{e(Y)} v_i \leq |X| \leq |Y| \sum_{i=0}^{\rho(Y)} v_i. \tag{62}$$

Thus, we have $e(Y) \leq \rho(Y)$. A code Y for which $e(Y) = \rho(Y)$ is called *perfect*. A code Y is perfect if and only if $d(Y) = 2\rho(Y) + 1$ or, equivalently, the spheres $S_{\rho(Y)}(y)$, $y \in Y$, form a partition of X.

From the existence of polynomials which are dual annihilators for codes we can derive some inequalities between their fundamental parameters.

Theorem 11: For any code Y in an n-class P-polynomial scheme X,

1. $d(Y) + d'(Y) \leq n + 2$.
2. $d(Y) \leq 2s'(Y) - \gamma'(Y) + 1$; equality implies $|Y|\Omega_P(f) = |X|$ where $f(t) = (1 + t)^{-\gamma'(Y)}(f^{Y,P}(t))^2$.
3. If $d(Y) \geq s'(Y)$, then Y is distance invariant and $a_i'(x, Y) = |X| g_0^{D'(Y),i,P}(P) - |Y| g^{D'(Y),i,P}(1)$, for any $x \in Y$ and $i \in D'(Y)$.
4. If σ_P is standard and $d(Y) \geq 2l - \theta + 1$, then

$$d'(Y) \leq d_{l-\theta}^{1,\theta}(P)$$

with equality if and only if $l = s'(Y), \theta = \gamma'(Y)$ and $(1 + t)^\theta P_{l-\theta}^{1,\theta}(t)$ is dual-minimal for Y.

A simple proof of Theorem 11 is based on the fact that for any $f \in F_n[t]$, (33) with $h = f(\sigma_P)$ and $a_i = a_i(x, Y)$, $x \in Y$, gives

$$|X| f_0(P) + |X| \sum_{i=d(Y)}^{n} a_i(x, Y) f_i(P)$$
$$= |Y| f(1) + \sum_{k=d'(Y)}^{n} a_k'(x, Y) f(\sigma_P(k)). \tag{63}$$

Therefore, if f is a dual annihilator for Y such that $f(1) = 0$ and $f_0(P) > 0$, then $d(Y) \leq \deg f$. The polynomials $f(t) = (1 - t) g^{d'(Y),P}(t)$ and $f(t) = (1 - t)\left(f^{Y,P}(t)\right)^2 / (1 + t)^{\gamma'(Y)}$ (see (45)) have these properties and give rise to the first and second statements. The third statement is obtained if one uses the polynomial $g^{D'(Y),i,U}$ of degree $s'(Y) - 1$ in (63) and takes (31) into account. To prove the last statement note that the left-hand side of (63) equals zero for $f(t) = (1 - t)(1 + t)^\theta \left(P_{l-\theta}^{1,\theta}(t)\right)^2 / (t_{l-\theta}^{1,\theta}(P) - t)$ since $f_0(P) = 0$ (see (48) and (49) for $a = 1, b = \theta$) and $d(Y) > \deg f$. Moreover, $s'(Y) \geq l$ by the second statement and $f(\sigma_P(k)) \geq 0$ if $k \geq d_{l-\theta}^{1,\theta}(P)$.

We can apply similar arguments to the rows of the outer distribution M which has rank $s'(Y)+1$ by Theorem 7. Considering in (33) $a_i = a_i(x,Y), x \in X$, and $h = f(\sigma_P)$ with a dual annihilator $f \in F_n[t]$ for Y we find that

$$|X|\sum_{i=0}^{n} a_i(x,Y)f_i(P) = |Y|f(1). \qquad (64)$$

In particular, for the dual minimal polynomial $f = f^{Y,P}$ we have $|X|\sum_{i=0}^{s'(Y)} a_i(x,Y)f_i(P) = |Y|$ and for any integer $k, s'(Y) < k \leq n$, and $f(t) = (1-t)^{k-s'(Y)}f^{Y,P}(t)$ we have $\sum_{i=0}^{k} a_i(x,Y)f_i(P) = 0$ with $f_k(P) \neq 0$. Since in both cases $f_i(P)$ depends only on $D'(Y)$, this allows one to compute the outer distribution of a code Y from a "small set of data".

Theorem 12: [37] Each column M_i of the outer distribution $M = [M_0, ..., M_n]$ of Y is a linear combination of the all-one vector $\mathbf{j}$ and the first $s'(Y)$ columns $M_0, ..., M_{s'(Y)-1}$, the coefficients of which are determined by the inner distribution.

Thus, the first $s'(Y)+1$ entries of any row $M(x)$, $x \in X$, of the outer distribution of a code Y are not all equal to zero and they uniquely determine the remaining entries of the row. This has some interesting consequences presented in [37]. In particular,

$$\rho(Y) \leq s'(Y)$$

and, hence, for any Y,

$$|Y|\sum_{i=0}^{s'(Y)} v_i \geq |X|. \qquad (65)$$

Moreover, if $d(Y) \geq 2s'(Y)-1$, then the first $s'(Y)$ entries of any row $M(x), x \in X$, are all zero except for $a_i(x,Y) = 1$ when $i = \partial(x,Y)$. For $i = s'(Y)$, the entry $a_i(x,Y)$ is uniquely determined from (64). In particular, this gives $a_i(x,Y) = |Y|/\left(|X|f_i^{Y,P}(P)\right)$ when $i = s'(Y) = \partial(x,Y)$ (see Example 1 below). Therefore, Y is completely regular if $d(Y) \geq 2s'(Y)-1$.

Note that from (33) it also follows that for any $f \in F_{\rho(Y)-1}[t]$ there exists $x \in X$ such that

$$|Y|f(1) + \sum_{k=d'(Y)}^{n} a'_k(x,Y)f(\sigma_P(k)) = 0.$$

This fact was used in [69], [83], [118] for obtaining asymptotic upper bounds for the covering radius $\rho(Y)$ of linear codes $Y \subseteq H_q^n$ when the dual distance $d'(Y)$ grows linearly with n. This approach is based on the inequalities $|a'_i(x,Y)| \leq a'_i(Y)$, $i \in N_n$, $x \in X$, which are satisfied by all linear codes $Y \subseteq H_q^n$; it makes use of the Chebyshev polynomials, characterized by the fact that they exhibit the smallest deviation from zero.

We now give the linear programming bounds which follow from solutions of the above extremum problems for the system P (see (46), (51), (59) and (60)).

Theorem 13: Let Y be any code in a P-polynomial scheme X. Then

$$|Y|K_P(d(Y)) \leq |X| \qquad (66)$$

with equality if and only if $d(Y) = 2s'(Y) - \gamma'(Y) + 1$ and $(t+1)^\theta P_{l-\theta}^{1,\theta}(t)$ is a dual minimal polynomial for Y where $l = s'(Y)$ and $\theta = \gamma'(Y)$.

In particular, for odd $d(Y) = 2l+1$ Theorem 13 gives the sphere packing bound (left-hand side of (62)) and implies that $P_l^{1,0}(t) = \sum_{i=0}^{l} P_i(t)$ is dual minimal for any perfect code Y. The latter is the (generalized) *Lloyd theorem* for perfect codes [17], [37].

Theorem 14: [80], [81] Let Y be a code in a P-polynomial scheme X with the standard function σ_P. If $d'(Y) \geq d_{l-\theta}^{1,\theta}(P)$, then

$$|Y|K_P(2l+1-\theta) \geq |X| \qquad (67)$$

with equality if and only if $d(Y) = 2s'(Y) - \gamma'(Y) + 1$ and $(t+1)^\theta P_{l-\theta}^{1,\theta}(t)$ is a dual minimal polynomial for Y where $l = s'(Y)$ and $\theta = \gamma'(Y)$.

Thus, the bounds (66) and (67) can be attained only simultaneously and in this case Y is a maximal d-code for $d = d(Y)$ and a minimal τ-design for $\tau = d'(Y) - 1$. Note that (65) gives a lower bound on the size of a code with a given number of dual "distances". Probably a stronger inequality

$$|Y|K_P(2s'(Y) - \gamma'(Y) + 1) \geq |X| \qquad (68)$$

holds which (together with (66)) would imply the statement 2 of Theorem 11. Then all bounds (66), (67), and (68) can be attained only simultaneously. In any case this is true for perfect codes (odd $d(Y)$). The inequality (68) was proved for the Hamming scheme in [78] but it is an open problem in the general case.

The following statement extends Theorem 14 from the case $d'(Y) = d_{l-\theta}^{1,\theta}(P)$ (see (56)) to the general case under the additional restriction that P satisfies the strengthened Krein condition. Therefore we do not repeat the necessary and sufficient conditions for this special case.

Theorem 15: [80], [81] Let X be a P-polynomial scheme with standard σ_P and assume that P satisfies the strengthened Krein condition. Then for any code $Y \subset X$,

$$|Y|L_P(\sigma_P(d'(Y))) \geq |X| \qquad (69)$$

with equality in the case $d'(Y) \neq d_{l-\theta}^{1,\theta}(P)$ if and only if $d(Y) = 2s'(Y) - \gamma'(Y)$ and the polynomial (57) with $U = P$ is dual minimal for Y where $\sigma = \sigma_P(d'(Y))$, $k = k_P(\sigma)$, $\varepsilon = \varepsilon_P(\sigma)$, $\gamma'(Y) = \varepsilon$, $s'(Y) = k + \varepsilon$.

Note that codes Y for which the bounds of Theorems 13, 14, 15 are attained belong to the class of codes satisfying $d(Y) \geq 2s'(Y) - \gamma'(Y)$ (cf. statement 2 of Theorem 11). There exists the following characterization of codes in this class.

Theorem 16: [77] Let Y be a code in a P-polynomial scheme X with the standard function σ_P such that $d =$

$d(Y) \geq 2$, $s' = s'(Y) \geq 1$, $\gamma' = \gamma'(Y)$, $\sigma = \sigma_P(d'(Y))$, $k = k_P(\sigma)$, $\varepsilon = \varepsilon_P(\sigma)$, and hence $t_{k-1+\varepsilon}^{1,1-\varepsilon}(P) \leq \sigma < t_k^{1,\varepsilon}(P)$. Then $d = 2s' - \gamma' + 1$ if and only if $s' = k$, $\gamma' = 1 - \varepsilon$, $\sigma = t_{k-1+\varepsilon}^{1,1-\varepsilon}(P)$, $|Y||K_P(d'(Y)) = |X|$, and $(t+1)^{1-\varepsilon} P_{k-1+\varepsilon}^{1,1-\varepsilon}(t)$ is dual minimal for Y; and $d = 2s' - \gamma'$ if and only if $s' = k + \varepsilon$, $\gamma' = \varepsilon$, $\sigma \neq t_{k-1+\varepsilon}^{1,1-\varepsilon}(P)$, $|Y||L_P(\sigma) = |X|$, and the polynomial (57) with $U = P$ is dual minimal for Y.

Note that for the class of codes Y in a P-polynomial scheme X defined by the condition $d(Y) \geq 2s'(Y) - \gamma'(Y) \geq 2$, the only parameter $|Y|$ (or $d'(Y)$) uniquely determines all fundamental parameters, the inner distribution, its Q-transform (or "dual distribution"), and the outer distribution of the code Y. Indeed, by Theorem 16 we know the dual minimal polynomial $f^{Y,P}(t)$ and hence the set $D'(Y) = \{i_1, ..., i_{s'(Y)}\}$ of integers which are dual distances, $d(Y)$, $s'(Y)$, $\gamma'(Y)$. From Theorem 10 and statement 3 of Theorem 11 it follows (by use of the polynomial $g^{D'(Y),i_j,P}(t)$ in (58)) that for any code Y in the class,

$$a'_{i_j}(Y) = |X|\rho_{k+\varepsilon-j}^{(\sigma)}(P), \quad j = 1, ..., s'(Y), \tag{70}$$

where $\sigma = \sigma_P(d'(Y))$, $k = k_P(\sigma)$, $\varepsilon = \varepsilon_P(\sigma)$. The inner distribution of Y and, in particular, the parameters $s(Y)$, $\gamma(Y)$ can be found with the help of (31). Moreover, all codes Y in this class are distance invariant and completely regular. This allows us to compute the outer distribution of Y with the help of (64) as was explained above. Note that from Theorem 16 it follows that the condition on a dual minimal polynomial in Theorems 13, 14, 15 is a consequence of the first condition and can be omitted.

Example 1 (continued) Apply these results to a code Y in the Hamming scheme H_2^{24} for which $d(Y) = 8$ or $d'(Y) = 8$. Since $d_3^{1,1}(P) = d_3(22) + 1 = 8$ and $K_P(8) = 2\sum_{i=0}^{3}\binom{23}{i} = 2^{12}$ (see (50) and (52)), we have $|Y| \leq 2^{12}$ or $|Y| \geq 2^{12}$ respectively by (66) and (67). Consider a code Y for which either of these bounds is attained. Any such Y is a maximal 8-code and a minimal 7-design, and must have the following properties: $d(Y) = 8$, $d'(Y) = 8$, $s'(Y) = 4$, $\gamma'(Y) = 1$, and $2^{12} f^{Y,P}(t) = \frac{385}{512}(t+1)P_3^{1,1}(t) = 2^8 t(t+1)(9t^2-1) = \sum_{i=0}^{3} P_i(t) + \frac{1}{6}P_4(t)$. Since $\sigma_P(d) = 1 - 2d/n$, $D'(Y) = \{8, 12, 16, 24\}$. Using the statement 3 of Theorem 11 (or (70)) and (31) we can find that $D(Y) = D'(Y)$, $a_8(Y) = a_{16}(Y) = 759$, $a_{12}(Y) = 2576$, $a_0(Y) = a_{24}(Y) = 1$, and $a'_i(Y) = 2^{12} a_i(Y)$ for all $i \in N_n$. (This means that Y must be *formally self-dual.*) Finally, $d(Y) \geq 2s'(Y) - 1$ and hence Y is completely distance regular. By use of the method explained above, we can mechanically compute the outer distribution M of the code Y. The following table gives the entries $M_i(x)$ for $0 \leq i \leq 8$, and for all $x \in \mathbb{F}_2^{24}$.

dist	0	1	2	3	4	5	6	7	8	#
0	1								759	1
1		1						253		24
2			1				77		352	276
3				1		21		168		2024
4					6		64		360	1771

In fact, a code Y having the above properties exists and is unique (within equivalence); it is the *extended binary Golay code* $Y = G_2(24, 12)$ ([88]). Thus, our table gives the *weight distribution of all cosets* of $G_2(24, 12)$.

Example 2 (continued) Apply (66) and (67) to a code Y in the Johnson scheme J_8^{24} for which $d(Y) = 4$ or $d'(Y) = 6$. Since $d_1^{1,1}(P) = 6$ and $K_P(4) = \sum_{i=0}^{1}\binom{7}{i}\binom{17}{i+1} = 969$ (see (53)), we have $|Y| \leq 759$ or $|Y| \geq 759$ respectively. Consider a code Y for which either of these bounds is attained. It must have the following properties: $d(Y) = 4$, $d'(Y) = 6$, $s'(Y) = 2$, $\gamma'(Y) = 1$, and $114 f^{Y,P}(t) = \frac{15}{19}(t+1)P_1^{1,1}(t) = (t+1)(23+34t) = \frac{2}{17}\left(P_0(t) + P_1(t) + \frac{1}{4}P_2(t)\right)$. Because $\sigma_P(d) = 1 - 2\frac{d(v+1-d)}{n(v+1-n)}$, $D'(Y) = \{6, 8\}$. Using the statement 3 of Theorem 11 and (31) we can find that $a'_6(Y) = 262752$, $a'_8(Y) = 471960$, $D(Y) = \{4, 6, 8\}$ and $a_4(Y) = 280$, $a_6(Y) = 448$, $a_8(Y) = 30$. Again in this case Y is completely distance regular, since $d(Y) \geq 2s'(Y) - 1$, and one can compute the outer distribution M of the code Y. In fact, a code Y having the above properties exists and is unique; it is the "octade code" formed by all vectors of Hamming weight 8 in the extended Golay code $G_2(24, 8)$ (see [88]).

More sophisticated applications can be found in [6], [36], [37], [76].

D. Codes and designs in Q-polynomial schemes

Let us consider codes in an n-class Q-polynomial scheme (X, R). Using (32) with $h = f(\sigma_Q)$ where $f \in F_n[t]$ are some annihilators for a code Y one can obtain the following dual analogue of Theorem 11.

Theorem 17: For any code Y in an n-class Q-polynomial scheme (X, R),

1. $d(Y) + d'(Y) \leq n + 2$.

2. $d'(Y) \leq 2s(Y) - \gamma(Y) + 1$; equality implies $|Y| = \Omega_Q(f)$ where $f(t) = (1+t)^{-\gamma(Y)}(f^{Y,Q}(t))^2$.

3. If $d'(Y) \geq s(Y)$, then Y is distance invariant and $a_i(x, Y) = |Y| g_0^{D(Y),i,Q}(Q) - g^{D(Y),i,Q}(1)$, for any $x \in Y$ and $i \in D(Y)$.

4. If σ_Q is standard and $d'(Y) \geq 2l - \theta + 1$, then

$$d(Y) \leq d_{l-\theta}^{1,\theta}(Q) \tag{71}$$

with equality if and only if $l = s(Y), \theta = \gamma(Y)$ and $(1+t)^\theta Q_{l-\theta}^{1,\theta}(t)$ is minimal for Y.

Many results concerning codes Y in Q-polynomial schemes are based on the existence of the representation

$$Q_k(\sigma_Q(\partial_R(x, y))) = \sum_{j=1}^{m_k} v_{k,j}(x)\overline{v_{k,j}(y)},\ k \in N_n \tag{72}$$

(see (10), (13)). In particular, let us emphasize the following important "dual" analogue of (65).

Theorem 18: (The absolute bound [37]) For any code Y in a Q-polynomial scheme,

$$|Y| \leq \sum_{j=0}^{s(Y)} m_j. \tag{73}$$

The proof of Theorem 18 is based on the fact that the $|Y|$ functions $f^{Y,Q}(\sigma_Q(\partial_R(x, y))$, $y \in Y$, belong to the space generated by the functions $v_{k,j}(x)$, $k \in N_{s(Y)}$, $j \in N^1_{m_k}$, equal to $\delta_{x,y}$ when $x \in Y$, and hence are linearly independent functions in $x \in X$.

In fact, the fundamental parameters of a code Y are determined from the sequence $a_i = a_i(Y)$, $i \in N_n$, and its Q-transform (see Section III-C). Analogously, for any $x \in X$, we can consider the sequence $a_i = a_i(x, Y)$, $i \in N_n$, and define the corresponding parameters; in particular, $d'(x, Y)$, $s(x, Y)$, $\gamma(x, Y)$. Note that, by (40), $d'(Y) \leq d'(x, Y)$ for any $x \in X$, and, similarly to statement 2 of Theorem 17, $d'(x, Y) \leq 2s(x, Y) - \gamma(x, Y)$ for any $x \in X \setminus Y$. Therefore, if we assume that $d'(Y) \geq 2l - \theta$ and $x \in X \setminus Y$, then $s(x, Y) \geq l$. For the polynomial $f(t) = (1+t)^\theta \left(Q^{0,\theta}_{l-\theta}(t)\right)^2 / (t^{0,\theta}_{l-\theta}(Q) - t)$ of degree $2l - \theta - 1$ we have $f_0(Q) = 0$ (see (48) and (49) for $a = 0, b = \theta$), and the use of $h = f(\sigma_Q)$ and $a_i = a_i(x, Y)$ in (32) shows that $\sum_{k=1}^n a_k(x, Y) f(\sigma_Q(k)) = 0$ under our assumption. If σ_Q is standard, then $f(\sigma_Q(k)) \geq 0$ if $d = d^{0,\theta}_{l-\theta}(Q) \leq k \leq n$ with equality in at most l points $k \in N^d_n$ while $s(x, Y) = |\{k \in N^1_n : a_k(x, Y) > 0\}| \geq l$.

This implies that for any code Y in a Q-polynomial scheme (X, R) with standard σ_Q such that $d'(Y) \geq 2l - \theta$,

$$\rho(Y) \leq d^{0,\theta}_{l-\theta}(Q) \tag{74}$$

with equality if and only if there exists a point $x \in X \setminus Y$ such that $(1+t)^\theta Q^{0,\theta}_{l-\theta}(t)$ is a polynomial of minimal degree which equals zero at $t = \sigma_Q(\partial_R(x, y))$ for any $y \in Y$.

The inequality (74) for the Hamming scheme is due to Tietäväinen [127]. For the general case it was given together with the necessary and sufficient condition for its attainability in [51]. Note that (71) and (74) give the following upper bounds on the packing and covering radius of a $(2l - \theta)$-design Y:

$$d(Y) \leq d^{1,\theta}_{l-\theta}(Q), \quad \rho(Y) \leq d^{0,1-\theta}_l(Q).$$

Let $R|Y$ be the set consisting of the $s(Y) + 1$ nonempty relations $R_i \cap Y^2$ (those with $a_i(Y) \neq 0$). Then $(Y, R|Y)$ is called the *restriction of* (X, R) *to* Y. It can be shown that if $d'(Y) \geq 2s(Y) - 1$, then $(Y, R|Y)$ is an $s(Y)$-class Q-polynomial scheme [37]. This theorem is a kind of "dual" of the result mentioned in Section IV-C about completely regular codes. The intersection numbers of this scheme can be computed with the help of the polynomials $f^{Y,Q}$ and $g^{D(Y),i,Q}$.

Next, we give linear programming bounds which follow from solutions of the above extremum problems for the system Q (see (46), (51), (59) and (60)). Recall that the representation (72) was used to prove that for the decomposable Q-polynomial schemes (in particular, for the Hamming and Johnson schemes) the system Q satisfies the strengthened Krein condition [77].

Theorem 19: [37], [47] For any code Y in a Q-polynomial scheme (X, R),

$$|Y| \geq K_Q(d'(Y)) \tag{75}$$

with equality if and only if $d'(Y) = 2s(Y) - \gamma(Y) + 1$ and $(t+1)^\theta Q^{1,\theta}_{l-\theta}(t)$ is a minimal polynomial for Y where $l = s(Y)$ and $\theta = \gamma(Y)$.

The well-known Rao bound [98] for τ-designs (orthogonal arrays) in the Hamming scheme and the Ray-Chaudhuri–Wilson bound [99] for τ-designs (block designs) in the Johnson scheme are special cases of (75) for these schemes (see (52) and (54)). A design which satisfies equality in (75) is called a *tight* τ*-design* where $\tau = d'(Y) - 1$. The subject of tight t-designs in the Johnson scheme has been introduced by Ray-Chaudhuri and Wilson [99] and has been investigated by several authors (see especially [7], [134], and the references therein). In particular, the polynomial $Q^{1,0}_l(t) = \sum_{i=0}^l Q_i(t)$ is minimal for any tight $2l$-design. This is (a generalized version of) the *Ray-Chaudhuri–Wilson theorem* for tight designs [99]; it is a "dual" of the Lloyd theorem for perfect codes [37], [101].

Theorem 20: [73], [77] Let Y be a code in a Q-polynomial scheme (X, R) with the standard function σ_Q. If $d(Y) \geq d^{1,\theta}_{l-\theta}(Q)$, then

$$|Y| \leq K_Q(2l + 1 - \theta) \tag{76}$$

with equality if and only if $d'(Y) = 2s(Y) - \gamma(Y) + 1$ and $(t+1)^\theta Q^{1,\theta}_{l-\theta}(t)$ is a minimal polynomial for Y where $l = s(Y)$ and $\theta = \gamma(Y)$.

Analogously, the bounds (75) and (76) can be attained only simultaneously and in this case Y is a minimal τ-design for $\tau = d'(Y) - 1$ and a maximal d-code for $d = d(Y)$. Note that the absolute bound (73) gives an upper bound on the size of a code with a given number of "distances". Probably a stronger inequality

$$|Y| \leq K_Q(2s(Y) - \gamma(Y) + 1) \tag{77}$$

holds which (together with (75)) would imply the statement 2 of Theorem 17. Then all bounds (75), (76), and (77) can be attained only simultaneously. In any case this is true for tight τ-designs with even τ (odd $d'(Y)$). The inequality (77) was proved in [77] for all decomposable Q-polynomial schemes (in particular, for the Hamming and Johnson schemes), but it is an open problem in the general case.

The following statement extends Theorem 20 from the case $d(Y) = d^{1,\theta}_{l-\theta}(Q)$ (see (56)) to the general case under the additional restriction that Q satisfies the strengthened Krein condition. Therefore we do not repeat the necessary and sufficient conditions for this special case.

Theorem 21: [73], [77] Let (X, R) be a Q-polynomial scheme with standard σ_Q and assume that Q satisfies the strengthened Krein condition. Then for any code Y in (X, R),

$$|Y| \leq L_Q(\sigma_Q(d(Y))) \tag{78}$$

with equality in the case $d(Y) \neq d^{1,\theta}_{l-\theta}(Q)$ if and only if $d'(Y) = 2s(Y) - \gamma(Y)$ and the polynomial (57) with $U = Q$ is minimal for Y where $\sigma = \sigma_Q(d(Y))$, $k = k_Q(\sigma)$, $\varepsilon = \varepsilon_Q(\sigma)$, $\gamma(Y) = \varepsilon$, $s(Y) = k + \varepsilon$.

For the Hamming and Johnson schemes the bound (78) in the first interval when $d^{1,0}_1(Q) \leq d(Y) \leq n$ coincides with the well-known Plotkin and Johnson bounds respectively (see [88] and the calculation of $L_Q(\sigma)$ in Section IV-B). For $d(Y) < d^{1,0}_1(Q)$, (78) improves upon these bounds. (We remind that (78) is true in the second interval and in all odd intervals without the restriction of the strengthened Krein condition.)

Note that codes Y for which the bounds of Theorems 19, 20, 21 are attained belong to the class of codes satisfying $d'(Y) \geq 2s(Y) - \gamma(Y)$ (cf. statement 2 of Theorem 17) and forming $s(Y)$-class Q-polynomial schemes. There exists the following characterization of codes in this class.

Theorem 22: [77] Let Y be a code in a Q-polynomial scheme (X, R) with the standard function σ_Q such that $d' = d'(Y) \geq 2$, $s = s(Y) \geq 1$, $\gamma = \gamma(Y)$, $\sigma = \sigma_Q(d(Y))$, $k = k_Q(\sigma)$, $\varepsilon = \varepsilon_Q(\sigma)$, and hence $t^{1,1-\varepsilon}_{k-1+\varepsilon}(Q) \leq \sigma < t^{1,\varepsilon}_k(Q)$. Then $d' = 2s - \gamma + 1$ if and only if $s = k$, $\gamma = 1 - \varepsilon$, $\sigma = t^{1,1-\varepsilon}_{k-1+\varepsilon}(Q)$, $|Y| = K_Q(d(Y))$, and $(t+1)^{1-\varepsilon} Q^{1,1-\varepsilon}_{k-1+\varepsilon}(t)$ is a minimal polynomial for Y; and $d' = 2s - \gamma$ if and only if $s = k + \varepsilon$, $\gamma = \varepsilon$, $\sigma \neq t^{1,1-\varepsilon}_{k-1+\varepsilon}(Q)$, $|Y| = |L_Q(\sigma)|$, and the polynomial (57) with $U = Q$ is minimal for Y.

For the class of codes Y in a Q-polynomial scheme (X, R) defined by the condition $d'(Y) \geq 2s(Y) - \gamma(Y) \geq 2$, the only parameter $|Y|$ (or $d(Y)$) uniquely determines all fundamental parameters, the inner and dual distributions of the code Y, and also the intersection numbers of the Q-polynomial scheme formed by Y. Indeed, by Theorem 22 we know the minimal polynomial $f^{Y,Q}(t)$ and hence the set $D(Y) = \{i_1, ..., i_{s(Y)}\}$ of integers which are "distances", $d'(Y)$, $s(Y)$, $\gamma(Y)$. From Theorem 10 and statement 3 of Theorem 17 it follows (by use of the polynomial $g^{D(Y), i_j, Q}(t)$ in (58)) that for any code Y in the class

$$a_{i_j}(Y) = |Y| \rho^{(\sigma)}_{k+\varepsilon-j}(Q), \quad j = 1, ..., s(Y),$$

where $\sigma = \sigma_Q(d(Y))$, $k = k_Q(\sigma)$, $\varepsilon = \varepsilon_Q(\sigma)$. The dual distribution of Y is computed by (30). Codes Y in the class are distance invariant and form $s(Y)$-class Q-polynomial schemes whose intersection numbers are determined with the help of the polynomials $f^{Y,Q}$ and $g^{D(Y),i,Q}$ (see Theorem 5.25 in [37] or Theorem 3.21 in [81]). Note that from Theorem 22 it follows that the condition on a minimal polynomial in Theorems 19, 20, 21 is a consequence of the first condition and can be omitted.

E. Bounds for spherical codes and designs

The results of Section IV-D are applicable to infinite distance-transitive spaces (see Section II-D). The unit sphere $S^{n-1} = \left\{x = (x_1, ..., x_n) \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 = 1\right\}$ is a distance-transitive space with the Euclidean distance $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ and the isometry group $G = O(n)$ consisting of all orthogonal matrices of order n. Any finite set $Y \subset S^{n-1}$ (called a *spherical code*) is characterized by the finite set $D(Y)$ of distinct nonzero values of $d(x, y)$ when $x, y \in Y$. It allows us to define for any spherical code Y the *minimum distance* $d(Y) := \min D(Y)$ and the *degree* $s(Y) := |D(Y)|$, and also the parameter $\gamma(Y)$ which equals 1, if the diameter $\sqrt{2}$ of S^{n-1} belongs to $D(Y)$, and equals zero otherwise. We can also measure the distance between $x, y \in S^{n-1}$ by the angle $\varphi = \varphi(x, y)$, $0 \leq \varphi \leq \pi$, where

$$\cos \varphi(x, y) = \sum_{i=1}^n x_i y_i = 1 - \frac{1}{2} d^2(x, y)$$

and denote by $\varphi(Y)$ the *minimum angular distance* between distinct points of Y. It is clear that $d(Y) = 2 \sin \frac{\varphi(Y)}{2}$. The normalized invariant measure μ on S^{n-1} is the normalized surface area of S^{n-1}. Let $\sigma_{n-1}(\varphi)$ be the surface area of $S^{n-1}(x, \varphi) = \{y : y \in S^{n-1}, \varphi(x, y) \leq \varphi\}$ and let $\sigma_{n-1} = 2\sigma_{n-1}(\frac{\pi}{2})$ be the surface area of S^{n-1}. It is well known that $\sigma_{n-1} = 2\pi^{\frac{n}{2}} / \Gamma(\frac{n}{2})$ and

$$\mu(S^{n-1}(x, \varphi)) = \frac{\sigma_{n-1}(\varphi)}{\sigma_{n-1}} = c_n \int_{\cos\varphi}^{1} (1 - z^2)^{\frac{n-3}{2}} dz \quad (79)$$

where $c_n = \Gamma(\frac{n}{2}) / \left(\Gamma(\frac{n-1}{2}) \Gamma(\frac{1}{2})\right)$ and $\Gamma(x)$ is the gamma function. The inner and outer distributions of a spherical code Y are given by the values

$$a_d(Y) := \frac{1}{|Y|} |\{(x, y) : x, y \in Y, d(x, y) = d\}|, \quad d \in [0, \sqrt{2}]$$

and, for any $x \in S^{n-1}$,

$$a_d(x, Y) := |\{(x, y) : y \in Y, d(x, y) = d\}|, \quad d \in [0, \sqrt{2}].$$

Note that $a_d(Y)$ and $a_d(x, Y)$ considered as functions of $d \in [0, \sqrt{2}]$, differ from zero only at a finite set of points d. For any function a_d which has this property (in particular, for $a_d(Y)$ and $a_d(x, Y)$) we can define its *Q-transform* as the infinite sequence $(a'_i)_{i=0}^{\infty}$ where

$$a'_i := \sum_{d \in [0, \sqrt{2}]} a_d q_i(d), \quad i = 0, 1, ...$$

As in the case of association schemes we can define (see Section III-C) the set $D'(Y)$, the *dual distance* $d'(Y)$ of Y and its *strength* $\tau(Y) = d'(Y) - 1$. Spherical τ-designs Y (i.e., $\tau(Y) \geq \tau$) were introduced in [45], in connection with an approximation formula for the evaluation of multidimensional integrals over S^{n-1} of the following form:

$$\int_{S^{n-1}} u(x) d\mu(x) \approx \frac{1}{|Y|} \sum_{x \in Y} u(x). \quad (80)$$

The code $Y \subset S^{n-1}$ is a *τ-design* if and only if the approximation formula (80) becomes equality for all functions $u(x)$ which are polynomials in coordinates of $x = (x_1, ..., x_n) \in S^{n-1}$ of degree at most τ. Thus $B(S^{n-1}, \tau+1)$ is the minimum number of nodes in the approximation formula under consideration.

Now we verify that S^{n-1} is "Q-polynomial" with respect to the standard change of variable $\sigma(d) = 1 - d^2/2$ (this means that $\sigma(d)$ is a decreasing continuous function such that $\sigma(0) = 1$, $\sigma(\sqrt{2}) = -1$). Indeed, from (28) and (79) it follows that the orthogonality and normalization conditions

$$c_n \int_{-1}^{1} Q_i(t) Q_j(t) (1 - t^2)^{\frac{n-3}{2}} dt = m_i \delta_{i,j}, \quad Q_k(1) = m_k, \quad (81)$$

uniquely define polynomials $Q_k(t)$ of degree k such that $q_k(z) = Q_k(\sigma(z))$ $(k = 0, 1, ...)$ and hence

$$\tilde{E}_k(x, y) = q_k(\partial(x, y)) = Q_k(\sigma(d(x, y))).$$

In the case of S^{n-1} the subspace V_k consists of all homogeneous harmonic polynomials in $x = (x_1, ..., x_n) \in S^{n-1}$ of degree k and has the dimension $m_k = \binom{n+k-1}{n-1} - \binom{n+k-3}{n-1}$, $k = 0, 1, \dots$. Thus

$$Q_i(t) = m_i P_i^{\frac{n-3}{2}, \frac{n-3}{2}}(t)$$

where

$$P_i^{\alpha,\beta}(t) = \frac{1}{2^i \binom{i+\alpha}{\alpha}} \sum_{j=0}^{i} \binom{i+\alpha}{j} \binom{i+\beta}{i-j} (t-1)^{i-j}(t+1)^j$$

are the *Jacobi polynomials* normalized by $P_i^{\alpha,\beta}(1) = 1$. (The Jacobi polynomials $P_i^{\alpha,\beta}(t)$ with $\alpha = \beta$ are called the *Gegenbauer polynomials*.) All results of Section IV-D are valid for spherical codes except for the statement 1 of Theorem 17 whose proof uses the finiteness of association schemes. (The absolute bound for S^{n-1} was proved in [45] in the strong form (77).) In particular, the $K_Q(d)$- and $L_Q(\sigma)$-problems and their solutions are valid for countable systems Q (compare (47) with (81)) and give rise to the following two results.

Theorem 23: (The DGS bound [45]) For a code $Y \subset S^{n-1}$, let $\tau(Y) = 2l - \theta$. Then

$$|Y| \geq \binom{l+n-2}{n-1} + \binom{l+n-1-\theta}{n-1} \tag{82}$$

with equality if and only if $\tau(Y) = 2s(Y) - \gamma(Y)$.

Note that if $p_i^{\alpha,\beta}$ is the largest zero of $P_i^{\alpha,\beta}(t)$ and $t_i^{a,b}$ is the largest zero of $Q_i^{a,b}(t)$, then $t_i^{a,b} = p_i^{\frac{n-3}{2}+a, \frac{n-3}{2}+b}$ and that Q satisfies the strengthened Krein condition.

Theorem 24: [73], [74] For a code $Y \subset S^{n-1}$, let $\sigma = \cos\varphi(Y)$, $\alpha = \frac{n-3}{2}$ and $k = k_Q(\sigma)$, $\varepsilon = \varepsilon_Q(\sigma)$. Then $|Y|$ does not exceed

$$\left(\binom{n+k-2}{n-1} + \binom{n+k-3+\varepsilon}{n-1}\right)\left(1 - \frac{P_{k-1}^{\alpha+1,\alpha+\varepsilon}(\sigma)}{P_k^{\alpha,\alpha+\varepsilon}(\sigma)}\right) \tag{83}$$

and, in particular, if $\sigma \leq p_{l-\theta}^{\alpha+1,\alpha+\theta}$, then

$$|Y| \leq \binom{l+n-2}{n-1} + \binom{l+n-1-\theta}{n-1}. \tag{84}$$

The bound (83) is attained if and only if

$$\tau(Y) \geq 2s(Y) - \gamma(Y) - 1 \geq 1.$$

The class of spherical codes for which the bounds of Theorems 23 and 24 are attained is described by Theorem 22; these codes carry Q-polynomial subschemes. Many examples can be found in [45], [74], [77]. In particular, the tight 7-design in S^7 containing 240 points and the tight 11-design in S^{23} containing 196560 points are the maximal codes with the angular distance $60°$; they allow one to determine the *kissing numbers* in dimensions 8 and 24 [74], [96].

The following asymptotic bound follows from (84). However, it was obtained earlier with the help of the MRRW polynomials (61).

Theorem 25: (The KL bound [65]) For any fixed φ, $0 < \varphi < \frac{\pi}{2}$, and $n \to \infty$,

$$\frac{1}{n}\log A\left(S^{n-1}, 2\sin\frac{\varphi}{2}\right) \lesssim \frac{1+\sin\varphi}{2\sin\varphi}\log\frac{1+\sin\varphi}{2\sin\varphi} - \frac{1-\sin\varphi}{2\sin\varphi}\log\frac{1-\sin\varphi}{2\sin\varphi}.$$

It should be noted that Theorems 23 and 24 give the best linear programming bounds in the class of polynomials of restricted degree. The necessary and sufficient conditions for optimality of $f^{(\sigma)}(t)$ for the $L_Q(\sigma)$-problem without this restriction were found in [24], [25], together with an improvement of (83) in some range when these conditions are not fulfilled. On the other hand, in [135] there was found a continuous function $f(t)$ having the properties $f(t) \geq 0$ for $-1 \leq t \leq 1$ and $f_i(Q) \leq 0$, $i = \tau, \tau+1, ...$; this yields an improvement of the DGS bound (82) for τ-designs if τ is sufficiency large.

Theorem 26: [135] For any spherical design $Y \subset S^{n-1}$,

$$|Y|\sigma_{n-1}\left(\arccos p_{\tau(Y)}^{\frac{n-1}{2}, \frac{n-1}{2}}\right) \geq \sigma_{n-1}. \tag{85}$$

In a certain sense the sphere packing bound $|Y|\sigma_{n-1}(\varphi(Y)/2) \leq \sigma_{n-1}$ and the bound (85) are analogues of the bounds (62) and (69) for P-polynomial schemes.

The projective spaces in n dimensions over $\mathbb{R}$, $\mathbb{C}$, and quaternions $\mathbb{H}$ $(n = 2, 3, \dots)$ also are Q-polynomial; the corresponding systems Q are systems of Jacobi polynomials and satisfy the strengthened Krein condition. Elements of these spaces can be considered as lines going through the origin. The results of Section IV-D are applicable to codes and designs in the projective spaces which were studied earlier in [44], [60], [65], [77]. The bounds for codes in the projective spaces have been successfully used to estimate "crosscorrelation" of codes [65], [75], [76], [109].

F. Universal bounds for codes and designs in P- and Q-polynomial schemes. Asymptotic results

In this subsection we consider a code Y in an n-class P- and Q-polynomial scheme X and tacitly suppose that the functions σ_P and σ_Q are standard. Of course, all results of Sections IV-C and IV-D are applicable. We give three pairs of *universal* bounds for codes and designs in such schemes. The term "universal" reflects the fact that these bounds are valid for all codes in all schemes under consideration.

First, for a P- and Q-polynomial association scheme X we extend the duality in bounding the optimal sizes of d-codes and $(d-1)$-designs to the polynomial case. For any $f \in F_n[t]$ and U (we use U for either P or Q, and use $\overline{U}$ for the other one), we define an U-dual polynomial $f^{(U)}$ to f as follows:

$$f^{(U)} := |X|^{-\frac{1}{2}} \sum_{i=0}^{n} f(\sigma_{\overline{U}}(i)) U_i$$

(cf. (39)). Analogously using (42)-(44) one can show that $f^{(U)}(\sigma_U(i)) = |X|^{\frac{1}{2}} f_i(\overline{U})$ and hence $f = (f^{(U)})^{(\overline{U})}$, $\Omega_{\overline{U}}(f)\Omega_U(f^{(U)}) = |X|$, and f has the property $\mathfrak{A}_{\overline{U}}(d)$ or $\mathfrak{B}_{\overline{U}}(d)$ if and only if $f^{(U)}$ has the property $\mathfrak{B}_U(d)$ or $\mathfrak{A}_U(d)$, respectively. In particular, for any $d \in N_n^1$, the following equalities hold [80]:

$$A_P(X,d)B_Q(X,d) = B_P(X,d)A_Q(X,d) = |X|. \tag{86}$$

As an example, consider the polynomial $g^{d,U}(t)$ defined by (45) and note that $g^{d,U}(\sigma_U(j)) \geq 0$ for $j \in N_n$ according to the assumption that σ_U is standard. Using the orthogonality condition and the property $U_k(1)\overline{U}_i(\sigma_{\overline{U}}(k)) = \overline{U}_i(1)U_k(\sigma_U(i))$ (see (42)-(44)) one can check that the $\overline{U}$-dual of $g^{d,U}(t)$ is the polynomial $|X|^{\frac{1}{2}} g^{n-d+2,\overline{U}}(t)/g_0^{d,U}(U)$ and hence $g^{d,U}(t)$ has the properties $\mathfrak{A}_U(d)$ and $\mathfrak{B}_U(n-d+2)$. This shows that

$$A_U(X,d) \leq 1/g_0^{d,U}(U) = |X| g_0^{n-d+2,\overline{U}}(\overline{U}),$$

$$B_U(X,d) \geq 1/g_0^{n-d+2,U}(U) = |X| g_0^{d,\overline{U}}(\overline{U}).$$

and implies the *first pair of universal bounds* [81]:

$$|X| g_0^{d'(Y),P}(P) \leq |Y| \leq 1/g_0^{d(Y),Q}(Q).$$

Each of these bounds is attained if and only if $d(Y) + d'(Y) = n+2$. In this case $g^{d(Y),Q}$ is an annihilator and $g^{d'(Y),P}$ is a dual annihilator for Y.

For the Hamming scheme H_q^n and the Johnson scheme J_n^v the first pair of universal bounds takes the following forms:

$$q^{d'(Y)-1} \leq |Y| \leq q^{n-d(Y)+1}, \tag{87}$$

$$\frac{\binom{v}{d'(Y)-1}}{\binom{n}{d'(Y)-1}} \leq |Y| \leq \frac{\binom{v}{n-d(Y)+1}}{\binom{n}{n-d(Y)+1}}, \tag{88}$$

respectively. These bounds for codes are called the Singleton and Johnson bounds [88], and codes satisfying equality in (87) or (88) are called MDS-codes and Steiner systems, respectively [88], [15]. In particular, (87) is attained for the Reed-Solomon codes and (88) is attained for the "octade" code (together with the four other bounds (66), (67), (75), (76)).

From the results of Sections IV-C and IV-D we deduce the *second pair of universal bounds* [37]:

$$K_Q(d'(Y)) \leq |Y| \leq \frac{|X|}{K_P(d(Y))} \tag{89}$$

with equality in the left- and right-hand side if and only if $d'(Y) = 2s(Y) - \gamma(Y) + 1$ and $d(Y) = 2s'(Y) - \gamma'(Y) + 1$, respectively.

Finally, if the systems Q and P satisfy the strengthened Krein condition, the results of Sections IV-C and IV-D imply the *third pair of universal bounds* [78], [81]:

$$\frac{|X|}{L_P(\sigma_P(d'(Y)))} \leq |Y| \leq L_Q(\sigma_Q(d(Y))) \tag{90}$$

with equality (when $d(Y) > 1$ and $d'(Y) > 1$) in the left- and right-hand side if and only if $d(Y) \geq 2s'(Y) - \gamma'(Y)$ or $d'(Y) \geq 2s(Y) - \gamma(Y)$, respectively.

The characterization of the codes for which the bounds in (89) and (90) are attained is given by Theorems 16 and 22. A list of the known codes in the Hamming and Johnson schemes for which (90) is attained can be found in [77], [78].

For finding asymptotic results the following special cases of bounds (89) and (90) are useful:

$$A_Q(X,d) \leq \begin{cases} |X|/\sum_{i=0}^k v_i & \text{if } d \geq 2k+1, \\ \sum_{i=0}^k m_i & \text{if } d \geq d_k^{1,0}(Q), \end{cases} \tag{91}$$

$$B_Q(X,d) \geq \begin{cases} \sum_{i=0}^k m_i & \text{if } d \geq 2k+1, \\ |X|/\sum_{i=0}^k v_i & \text{if } d \geq d_k^{1,0}(P). \end{cases} \tag{92}$$

In particular, in the case of the Hamming scheme H_q^n, if $1 \leq k = k(n) \leq n(q-1)/q$ and $n \to \infty$, then $d_k^{1,0}(Q)/n = \gamma_q(k/n) + o(1)$, where

$$\gamma_q(x) = \tfrac{1}{q}\left(q - 1 - (q-2)x - 2\sqrt{(q-1)x(1-x)}\right)$$

(see [90] and [78]). Notice that $\gamma_q(x)$ is a decreasing continuous function on $[0,(q-1)/q]$ which coincides with its inverse function, that is, $\gamma_q(\gamma_q(x)) = x$. Therefore, if q is fixed and $\lim_{n\to\infty} \frac{d}{n} = \delta$, $0 \leq \delta \leq (q-1)/q$, then the second bounds in (91) and (92) give rise to the *first form of the MRRW bound for codes* [90]:

$$\frac{1}{n}\log_q A(H_q^n, d) \lesssim H(\gamma_q(\delta), q) \tag{93}$$

and the following asymptotic bound for designs [78]:

$$\frac{1}{n}\log_q B(H_q^n, d) \gtrsim 1 - H(\gamma_q(\delta), q), \tag{94}$$

where $H(x,q) = -x\log_q x - (1-x)\log_q(1-x) + \log_q(q-1)$ is the *Shannon entropy*. In the case of the Johnson scheme J_n^v, if $\lim_{n\to\infty} \frac{k}{v} = \zeta$ and $\lim_{v\to\infty} \frac{n}{v} = \eta$, where $0 \leq \zeta \leq \eta$ and $0 < \eta \leq \frac{1}{2}$, then $d_k^{1,0}(Q)/v = \xi_\eta(\zeta) + o(1)$, where, for any η, $0 \leq \eta \leq \frac{1}{2}$,

$$\xi_\eta(x) = \frac{\eta(1-\eta) - x(1-x)}{1 + 2\sqrt{x(1-x)}}$$

is a decreasing continuous function which maps the interval $[0,\eta]$ onto $[0,\eta(1-\eta)]$ (see [90] and [81]). The inverse function $\xi_\eta^{-1}(x)$ can be expressed in the following explicit form:

$$\xi_\eta^{-1}(x) = \frac{1}{2}\left(1 - \sqrt{1 - 4\left(\sqrt{\eta(1-\eta) - x(1-x)} - x\right)^2}\right).$$

Therefore, if $\lim_{v\to\infty} \frac{n}{v} = \eta$ and $\lim_{v\to\infty} \frac{d}{v} = \delta$, where $0 < \eta \leq \frac{1}{2}$ and $0 \leq \delta \leq \eta(1-\eta)$, then the second bound in (91) gives rise to the asymptotic bound [90]

$$\frac{1}{v}\log_2 A(J_n^v, d) \lesssim H\left(\xi_\eta^{-1}(\delta)\right) \tag{95}$$

where $H(x) = H(x, 2)$. On the other hand, as it was shown in [81], if $\lim_{v\to\infty} \frac{n}{v} = \eta$ and $\lim_{v\to\infty} \frac{k}{v} = \zeta$, where $0 \le \zeta \le \eta(1-\eta)$ and $0 < \eta \le \frac{1}{2}$, then $d_k^{1,0}(P)/v = \xi_\eta^{-1}(\zeta)+o(1)$, and hence, for $\lim_{v\to\infty} \frac{n}{v} = \eta$ and $\lim_{v\to\infty} \frac{d}{v} = \delta$, where $0 < \eta \le \frac{1}{2}$ and $0 \le \delta \le \eta$, the second bound in (92) gives rise to the asymptotic bound

$$\frac{1}{n}\log_2 B(J_n^v, d) \gtrsim H(\eta)-\eta H\left(\frac{\xi_\eta(\delta)}{\eta}\right)-(1-\eta)H\left(\frac{\xi_\eta(\delta)}{1-\eta}\right).$$

In the typical situation the second bounds in (91) and (92) (which follow from the third pair of universal bounds) are better than the first ones when the parameter d is sufficiently large and become worse when d is small. In particular, this is true for the bound (93) which for sufficiently small d is worse than the Hamming asymptotic bound $\frac{1}{n}\log_q A(H_q^n, d) \lesssim 1 - H(\delta, q)$. This raises the problem of "smoothing" these bounds. A similar problem of bounding the *Shannon reliability function* for probabilistic channels was considered in [108] where the *straight-line bound* was found. The *principle of the multiple packing* (applicable to the translation schemes, see Section V) gives the Bassalygo–Elias inequality [12]

$$\binom{n}{w} A(H_2^n, 2d) \le 2^n A(J_w^n, d) \tag{96}$$

where $1 \le w \le n/2$, and the straight-line bound for H_q^n [69]

$$A(H_q^n, d)\sum_{i=0}^{r}\binom{l}{i}(q-1)^i \le q^l A(H_q^{n-l}, d-2r)$$

where $r \le l, 2r \le d$. A consequence of (95) and (96) for $\lim_{n\to\infty} \frac{d}{n} = \delta$, $0 \le \delta \le \frac{1}{2}$, is the *second form of the MRRW bound for codes* [90]:

$$\frac{1}{n}\log_2 A(H_2^n, d) \lesssim 1 - \max_\eta \left(H(\eta) - H\left(\xi_\eta^{-1}\left(\frac{\delta}{2}\right)\right)\right),$$

where $\eta \in [\frac{1}{2}(1-\sqrt{1-2\delta}), \frac{1}{2}]$ and $\xi_\eta^{-1}(x)$ is defined above. This becomes better than (95) with $q = 2$ when $\delta < 0.272$.

The argument that leads to (96) does not apply to the minimal design problem. However, a similar result is valid in terms of the linear programming bounds. Rodemich (see [100], [42]) used the fact that any nonnegative definite function $h(\partial_H(x, y))$ on H_2^n is nonnegative definite on (subset) J_w^n as well and proved the following analogue of (96) for objective functions of linear programming bounds:

$$\binom{n}{w} A_Q(H_2^n, 2d) \le 2^n A_Q(J_w^n, d)\ . \tag{97}$$

Combination of (97) with (86) for $X = H_2^n$ and $X = J_w^n$ gives the following results [80]:

$$B_P(H_2^n, 2d) \ge B_P(J_w^n, d), \tag{98}$$

$$B(H_2^n, 2d) \ge B_P(J_w^n, d) = \frac{\binom{n}{w}}{A_Q(J_w^n, d)}. \tag{99}$$

In particular, (99) and (95) (considered as a bound on $A_Q(J_w^n, d)$) for $\lim_{n\to\infty} \frac{d}{n} = \delta$, $0 \le \delta \le \frac{1}{2}$, give the following asymptotic bound [80]:

$$\frac{1}{n}\log_2 B(H_2^n, d) \gtrsim \max_\eta \left(H(\eta) - H\left(\xi_\eta^{-1}\left(\frac{\delta}{2}\right)\right)\right), \tag{100}$$

where $\eta \in [\frac{1}{2}(1-\sqrt{1-2\delta}), \frac{1}{2}]$.

The essential difficulty in extending the second form of the MRRW bound to the case $q \ge 3$ is caused by the fact that the natural generalization of (96) connects H_q^n with subschemes of the association scheme considered in Example 3 which are not Q-polynomial. Some results in this direction were obtained in [1], [125].

Finally, in addition to the (nonconstructive) results (96)-(99) which give bounds for the Hamming space from bounds for the Johnson space, let us point out a "constructive" relationship between codes and designs in the Hamming and Johnson schemes. It is the celebrated *Assmus–Mattson theorem* [4], which allows one to obtain good combinatorial designs and constant weight codes from certain codes in the binary Hamming scheme. A strengthening of this result can be found in [29].

V. Translation schemes

A. Definitions and preliminaries

Certain association schemes are invariant under "translations", of the form $(x, y) \mapsto (x+z, y+z)$. Examples are the Hamming scheme and the composition scheme, described in Section II-A (Examples 1 and 3). We shall examine the appropriate generalization, under the name of "translation schemes", borrowed from [26]. A comprehensive treatment of that subject is given by Camion [32]. The material of this section is mainly taken from [42].

Definition 10: Let $(X, +)$ be a finite Abelian group, and let (X, R) be an n-class association scheme. Assume that (X, R) is $(X, +)$-*invariant*, i.e.,

$$\text{if } (x, y) \in R_i, \text{ then } (x+z, y+z) \in R_i,$$

for all $z \in X, i \in N_n$. Then (X, R) is said to be a *translation scheme with respect to the group* $(X, +)$.

We briefly examine the Hamming scheme H_q^n from this viewpoint. Let $H_1, \cdots, H_n$ be n Abelian groups of order q, and consider their direct product $X := H_1 \times \cdots \times H_n$. It is clear that H_q^n is a translation scheme with respect to *any* of these group structures.

This simple example shows that a given association scheme (X, R) may be a translation scheme with respect to several group structures. Further explanation of the phenomenon will be given in Section V-C.

For a translation scheme, the p-numbers and the q-numbers (Definition 3) can be determined as follows.

The homomorphisms of the group $(X, +)$ to the group $(\mathbb{C}^*, \cdot)$ are referred to as the (irreducible) *characters* of X. (Henceforth we often write X instead of $(X, +)$.) The set of characters of X has the structure of an Abelian group, isomorphic to X itself; this character group is called the *dual* of X and is denoted by X'. We shall use a bracket notation for characters, that is $\langle x, x'\rangle := x'(x)$ for $x \in X$ and $x' \in X'$. The group characters satisfy the *orthogonality relations*

$$\sum_{x\in X}\langle x, x'\rangle = |X|\delta_{0,x'}, \quad \sum_{x'\in X'}\langle x, x'\rangle = |X|\delta_{0,x}.$$

It is possible to identify X' with X so as to have the symmetry property $\langle x, x'\rangle = \langle x', x\rangle$ for all $x \in X, x' \in X$.

From the n-class translation scheme (X, R) we define a partition $\Pi := \{X_0, X_1, \cdots, X_n\}$ of the group X into $n+1$ blocks $X_i := \{x \in X : (x, 0) \in R_i\}$. This implies $X_0 = \{0\}$ and $X_{i^\sigma} = -X_i$ (for the pairing $i \mapsto i^\sigma$). The relation R_i can be recovered from the block X_i as follows:

$$R_i = \{(x, y) \in X^2 \,:\, x - y \in X_i\}. \tag{101}$$

It can be shown that there exists a unique partition $\Pi' = \{X'_0, X'_1, \cdots, X'_n\}$ of the dual group X' into $n+1$ blocks X'_k, with $X'_0 = \{0\}$ and with the following property. For $i, k \in N_n$, define $\phi_i : X' \to \mathbb{C}$ and $\psi_k : X \to \mathbb{C}$ by

$$\phi_i(x') := \sum_{x\in X_i}\langle x, x'\rangle, \quad \psi_k(x) := \sum_{x'\in X'_k}\langle x, x'\rangle. \tag{102}$$

Then ϕ_i is constant over each block X'_k and ψ_k is constant over each block X_i. More precisely,

$$\phi_i(x') = \overline{p_i(k)} \text{ for } x' \in X'_k, \quad k, i \in N_n, \tag{103}$$
$$\psi_k(x) = q_k(i) \text{ for } x \in X_i, \quad i, k \in N_n, \tag{104}$$

where the numbers $p_i(k)$ and $q_k(i)$ are the p-numbers and the q-numbers of (X, R).

The partition Π' (of X') will be called the *dual* of the partition Π (of X). From Π' we define a partition R' of X'^2 like in (101). More precisely, $R' := \{R'_0, R'_1, \cdots, R'_n\}$ with

$$R'_k := \{(x', y') \in X'^2 \,:\, y' - x' \in X'_k\}.$$

It can be shown that (X', R') is a translation scheme with respect to the dual group $(X', +)$. From (102) it follows that *the p-numbers of* (X', R') *are the q-numbers of* (X, R) *and conversely.* Thus, with an obvious notation, we have the duality relations $p'_k(i) = q_k(i)$ and $q'_i(k) = p_i(k)$, given in (25). In particular, $v'_k = m_k$ and $m'_i = v_i$.

In the context of the Krein duality (Section II-C), this leads us to introduce the following definition of duality for translation schemes. It is equivalent to a concept introduced by Tamaschke for commutative Schur rings [123].

Definition 11: Let (X, R) be a translation scheme, with respect to a given Abelian group $(X, +)$, and let $\Pi = \{X_i : i \in N_n\}$ be the corresponding partition of X. The translation scheme (X', R'), with respect to the dual group $(X', +)$ and corresponding to the dual partition $\Pi' = \{X'_k : k \in N_n\}$ of X', is called the *dual* of the translation scheme (X, R).

An interesting treatment of duality in translation schemes is given by Godsil, in relation with the theory of *equitable partitions* [54].

As an example, consider once again the Hamming scheme (X, R), together with the Abelian group structure $X = H_1 \times \cdots \times H_n$. In this case, the block X_i in Π consists of the q-ary n-tuples of weight i (i.e., those having i nonzero components). Let us identify the dual group X' with X. It turns out that the dual partition Π' coincides with the weight partition Π. Thus, for a suitable ordering, we have $X'_i = X_i$ for all $i \in N_n$, which shows that the Hamming scheme is *self-dual.* The formulas (103) and (104) lead to the expressions (17) of the p-and q-numbers in terms of the Krawtchouk polynomials.

Similarly, the composition scheme (Example 3 in Section II-A) is a self-dual translation scheme. There are other interesting families of that type [41], [43], [89]. Some examples of translation schemes that are *not* self-dual can be found in [30], [37].

B. Additive codes in a translation scheme

As a generalization of the classical notion of a *linear code* over a field alphabet (in Hamming scheme), we shall examine additive codes in a translation scheme.

Definition 12: A code Y in a translation scheme (X, R) is said to be *additive* if Y a subgroup of the underlying Abelian group $(X, +)$.

Consider the inner distribution $(a_0, a_1, \cdots, a_n)$ of an additive code Y (see Definition 5). It follows from (101) that a_i counts the code points belonging to the block X_i in the partition Π, that is

$$a_i = a_i(Y) = |Y \cap X_i|, \quad \text{for } i \in N_n.$$

Next, we define the "annihilator code" of an additive code Y, by generalizing the usual notion of the orthogonal code of a linear code in Hamming scheme. (The terminology is not standard, but "annihilator" seems preferable to "orthogonal", in the general setting.)

Definition 13: Let Y be an additive code in a translation scheme (X, R). The *annihilator code* of Y (with respect to the given group $(X, +)$) is the code Y° in the dual translation scheme (X', R') defined by

$$Y^\circ := \{x' \in X' \,:\, \langle x, x'\rangle = 1 \text{ for all } x \in Y\}.$$

It is clear that Y° is an additive code in (X', R'). Similarly, for an additive code V in (X', R'), we define its annihilator code to be ${}^\circ V := \{x \in X : \langle x, x'\rangle = 1 \text{ for all } x' \in V\}$. This is an additive code in (X, R). For double annihilators, we simply have

$$Y = {}^\circ(Y^\circ), \quad V = ({}^\circ V)^\circ.$$

The character group Y' of Y is related to Y° by $Y' = X'/Y^\circ$ (the group of coset codes $x' + Y^\circ$ with $x' \in X'$). This implies $|Y||Y^\circ| = |X|$. As a consequence of the orthogonality relations on group characters, we obtain

$$\sum_{x\in Y}\langle x, x'\rangle = \begin{cases} |Y| & \text{if } x' \in Y^\circ \\ 0 & \text{if } x' \notin Y^\circ \end{cases} \tag{105}$$

If Y is a linear code of length n over $\mathbb{F}_q$ (in the usual sense), then $Y^\circ = Y^\perp$, the orthogonal of Y. This stems from the fact that the characters of $(\mathbb{F}_q^n, +)$ are given by

$$\langle x, x' \rangle = \exp\left(\frac{2\pi i}{p}\mathrm{Tr}(x' x^T)\right),$$

for all $x, x' \in \mathbb{F}_q^n$. Here, Tr denotes the *trace* from the field $\mathbb{F}_q$ to its prime subfield $\mathbb{F}_p$. In the binary case, $q = p = 2$, this reduces to $\langle x, x' \rangle = (-1)^{x' x^T}$.

The next result is a generalization of the MacWilliams identities on the weight distributions of a linear code and its orthogonal. It produces a clear interpretation of Theorem 3 in the restricted framework of additive codes in translation schemes. The proof is based on (103), (104) and (105).

Theorem 27: (Generalized MacWilliams identities [37], [42]) The inner distribution $(a_k(Y^\circ) = |Y^\circ \cap X'_k|)_{k=0}^n$ of Y° is proportional to the Q-transform of the inner distribution $(a_i(Y) = |Y \cap X_i|)_{i=0}^n$ of Y. More precisely,

$$a_k(Y^\circ) = |Y|^{-1} a'_k(Y), \quad a_i(Y) = |Y^\circ|^{-1} a'_i(Y^\circ).$$

As a consequence, the fundamental parameters (see Section III-C) of a code Y in a translation scheme are related to those of its annihilator code Y° by $d(Y) = d'(Y^\circ)$, $d'(Y) = d(Y^\circ)$, $s(Y) = s'(Y^\circ)$, $s'(Y) = s(Y^\circ)$, $\gamma(Y) = \gamma'(Y^\circ)$, $\gamma'(Y) = \gamma(Y^\circ)$.

Finally, let us examine the outer distribution M of an additive code Y (see Definition 8). In view of (101), noting that $Y = -Y$, we obtain

$$M_i(x) = |(x + Y) \cap X_i|, \quad \text{for } i \in N_n, x \in X.$$

This means that the x-row $M(x)$ of M is the distribution of the *coset code* $x + Y$ with respect to the partition Π.

For the outer distribution M' of the annihilator code Y°, we similarly have $M'_k(x') = |(-x' + Y^\circ) \cap X'_k|$, with $k \in N_n$ and $x' \in X'$. As an extension of Theorem 27 we obtain the following expression for the Q-transform of the outer distribution rows:

$$\sum_{k=0}^{n} M'_k(x') p_i(k) = |Y^\circ| \sum_{y \in Y \cap X_i} \langle y, x' \rangle. \tag{106}$$

By use of (106) we can derive a remarkable result (Theorem 28 below) that allows us to decide whether a given additive code Y carries a (translation) subscheme of (X, R), and to characterize the dual of this subscheme. Note that by Theorem 7 and Theorem 27 the rank of the outer distribution M' of Y° is equal to $s(Y) + 1$.

Definition 14: Let Y be an additive code of degree $s = s(Y)$ in a translation scheme (X, R). If the restriction $(Y, R|Y)$ is an association scheme (with s classes), then it is called a *subscheme* of (X, R).

Theorem 28: [37] Given an additive code Y of degree s, the restriction $(Y, R|Y)$ is a subscheme of (X, R) if and only if the outer distribution M' of the annihilator code Y° has $s + 1$ distinct rows.

In this case, the dual scheme of the translation scheme $(Y, R|Y)$ is (Y', R^*) where $R^* = \{R_0^*, \cdots, R_s^*\}$ consists of the $s+1$ relations on $Y' := X'/Y^\circ$ defined as follows: a pair of coset codes $(x'_1 + Y^\circ, x'_2 + Y^\circ)$ belongs to a given relation R_k^* if and only if the outer distribution row $M'(x'_1 - x'_2)$ is a fixed $(n+1)$-tuple (among the $s+1$ possibilities).

For example, Theorem 28 applies to the extended binary Golay code Y examined in Section IV-C. Recall that Y is self-orthogonal: $Y = Y^\perp = Y^\circ$. The code Y has degree $s = 4$, and the outer distribution $M' = M$ of its orthogonal Y° has 5 $(= s + 1)$ distinct rows. The 4-class association scheme carried by Y is mentioned at the end of Section IV-D. Since $(Y, R|Y)$ is Q-polynomial, its dual scheme (Y', R^*), carried by the factor group $Y' = X'/Y^\circ$, is P-polynomial. It is the "distance scheme" for the cosets of the extended binary Golay code (see [26], p. 361).

The reader familiar with the Golay code may be interested in a more sophisticated example. Take Y to be the perfect binary Golay code of length 23 (and dimension 12). This code has degree $s = 7$. The orthogonal code $Y^\perp$ can be shown to be completely regular; its outer distribution M' has 8 distinct rows $M'(x)$, corresponding to the 8 values $\partial_H(x, Y^\perp) = 0, 1, \cdots, 6, 7$ (see [26], p. 362).

C. $\mathbb{Z}_4$-*additive binary codes*

Important research work has been devoted recently to the class of binary codes that are additive over $\mathbb{Z}_4$, the cyclic group of order 4 (see especially [57], [92]). This subsection aims at showing how that subject fits into the framework of association scheme theory.

From a group theoretic viewpoint, translation schemes can be presented as follows. Let $\mathrm{Aut}(X, R)$ denote the automorphism group of a given association scheme (X, R). Assume that $\mathrm{Aut}(X, R)$ contains an Abelian subgroup G which is *regular* on X, in the sense that G is transitive on X and has order $|X|$. This provides the point set X with the structure of an Abelian group $(X, +)$, isomorphic to G, through the definition

$$x_0^g + x_0^h := x_0^{gh}, \quad \text{for all } g, h \in G,$$

where x_0 is a fixed point in X. It is clear that (X, R) is a translation scheme with respect to $(X, +)$ (see Definition 10). In fact, the "translation structures" of (X, R) correspond exactly to the regular Abelian subgroups of $\mathrm{Aut}(X, R)$.

Consider the binary Hamming scheme H_2^n (see Example 2 in Section II-D). Its automorphism group is the monomial group (or hyperoctahedral group) M_n of degree n (and order $n!2^n$). As we shall see, M_n contains several regular Abelian subgroups.

For $n = 2$, the monomial group M_2, of order 8, is the symmetry group of the square $\{1, -1\}^2$. It contains an element g of order 4 that cyclically permutes the 4 vertices $(\pm 1, \pm 1)$, namely

$$g := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

In effect, for $p_0 := (1, 1)$ we obtain $p_1 := p_0 g = (1, -1)$, $p_2 := p_1 g = (-1, -1)$, $p_3 := p_2 g = (-1, 1)$, $p_0 = p_3 g$.

In terms of the usual binary alphabet $\{0,1\}$, this induces the cyclic permutation (x_0, x_1, x_2, x_3) of the binary ordered pairs, with

$$x_0 := (0,0),\ x_1 := (0,1),\ x_2 := (1,1),\ x_3 := (1,0). \quad (107)$$

As a conclusion, M_2 contains *two* regular Abelian subgroups: not only the elementary Abelian group $\mathbb{Z}_2 \times \mathbb{Z}_2$ (consisting of the diagonal matrices), but also the cyclic group $\mathbb{Z}_4$ generated by g.

Note that the cyclic permutation in (107) corresponds to the *Gray map* between $\mathbb{Z}_2 \times \mathbb{Z}_2$ and $\mathbb{Z}_4$, that is $(0,0) \mapsto 0$, $(0,1) \mapsto 1$, $(1,1) \mapsto 2$, $(1,0) \mapsto 3$. This map underlies the "concrete approach" to $\mathbb{Z}_4$-additive binary codes [57].

Let us now turn to the general case of H_2^n with $n \geq 2$. By considering partitions of the n coordinate positions in blocks of size 1 or 2, we obtain a whole class of regular Abelian subgroups of M_n, of the form

$$G = \mathbb{Z}_2^k \times \mathbb{Z}_4^m, \quad \text{with } k + 2m = n.$$

Each of the groups G (together with a corresponding coordinate partition) provides H_2^n with a translation scheme structure, and gives rise to a well-defined class of additive (binary) codes (see Definition 12).

Let Y be an additive code with respect to G, and let Y° be the annihilator code of Y (see Definition 13). In view of Theorem 27, the inner distributions of Y and Y° (which are their ordinary weight distributions) are related to each other by the MacWilliams identities in the *usual sense*.

The "homogeneous cases" are $G = \mathbb{Z}_2^n$, which yields the class of linear binary codes, and $G = \mathbb{Z}_4^{n/2}$ (for even n), which yields the class of $\mathbb{Z}_4$-*additive binary codes* [57]. In the latter case, the annihilator code Y° of Y is the natural orthogonal code $Y^\perp$ over the cyclic group $\mathbb{Z}_4$.

A very interesting example is provided by the *Kerdock codes* $\mathcal{K}$ and their $\mathbb{Z}_4$-orthogonals $\mathcal{K}^\perp$ which are the "Preparata codes" $\mathcal{P}$ (see [57]). Quotes are used here because $\mathcal{P}$ is not exactly the same as the official Preparata code, although they both have the same essential properties and, in particular, the same distance distribution. This example is quite remarkable for the following reason. It has been known for a long time that the weight distributions of the Kerdock and Preparata codes are the MacWilliams transform of each other, although these codes are *nonlinear* (over $\mathbb{F}_2$). The result in [57] alluded to above says that $\mathcal{K}$ is $\mathbb{Z}_4$-additive, and it identifies the $\mathbb{Z}_4$-orthogonal $\mathcal{K}^\perp$ of $\mathcal{K}$ as a certain "Preparata code" $\mathcal{P}$.

References

[1] M. Aaltonen, "A new upper bound on nonbinary block codes", *Discr. Math.*, vol. 83, pp. 139–160, 1990.

[2] R. Askey and J. Wilson, "A set of orthogonal polynomials that generalize the Racah coefficients or $6-j$ symbols", *SIAM J. Math. Anal.*, vol. 10, pp. 1008–1016, 1979.

[3] E.F. Assmus, Jr. and J.D. Key, *Designs and Their Codes.* Cambridge: Cambridge University Press, 1992.

[4] E.F. Assmus, Jr. and H.F. Mattson, Jr., "New 5-designs", *J. Combin. Theory*, vol. 6, pp. 122–151, 1969.

[5] E.F. Assmus, Jr. and H.F. Mattson, Jr., "Coding and combinatorics", *SIAM Review*, vol. 16, pp. 349–388, 1974.

[6] E.F. Assmus, Jr. and V. Pless, "On the covering radius of extremal self-dual codes", *IEEE Trans. Inform. Theory*, vol. 29, pp. 359–363, 1983.

[7] E. Bannai, "On tight designs", *Quart. J. Math. Oxford*, vol. 28, pp. 433–448, 1977.

[8] E. Bannai, "Tannaka–Krein duality for association schemes", *Linear Algebra Appl.*, vol. 46, pp. 83–90, 1982.

[9] E. Bannai, "Orthogonal polynomials in coding theory and algebraic combinatorics", in *Orthogonal Polynomials*, P. Nevai, Ed., Kluwer Acad. Publ., pp. 25–53, 1990.

[10] E. Bannai and T. Ito, *Algebraic Combinatorics I. Association Schemes.* Menlo Park, CA: Benjamin/Cummings, 1984.

[11] E. Bannai and T. Ito, "Current research on algebraic combinatorics", *Graphs and Combinatorics*, vol. 2, pp. 287–308, 1986.

[12] L.A. Bassalygo, "New upper bounds for error correcting codes", *Probl. Inform. Transm.*, vol. 1, no. 4, pp. 32–35, 1965.

[13] V. Belevitch, "Conference networks and Hadamard matrices", *Ann. Soc. Scient. Bruxelles*, vol. 82, pp. 13–32, 1968.

[14] M.R. Best and A.E. Brouwer, "The triply shortened binary Hamming code is optimal", *Discr. Math.*, vol. 17, pp. 235–245, 1977.

[15] Th. Beth, D. Jungnickel, and H. Lenz, *Design Theory.* Manheim: Bibl. Institut-Wissenschaftsverlag, 1985.

[16] J. Bierbrauer, K. Gopalakrishnan, and D.R. Stinson, "Bounds for resilient functions and orthogonal arrays", in *Advances in Cryptology–CRYPTO'94, Lecture Notes in Computer Science No. 839*, Y.G. Desmedt, Ed. New York: Springer-Verlag, 1994, pp. 247–256.

[17] N.L. Biggs, "Perfect codes in graphs", *J. Combin. Theory Ser. B*, vol. 15, pp. 289–296, 1973.

[18] N.L. Biggs, *Algebraic Graph Theory.* Cambridge: Cambridge University Press, 1974.

[19] H.F. Blichfeldt, "The minimum value of quadratic forms and the closest packing of spheres", *Math. Ann.*, vol. 101, pp. 605–608, 1929.

[20] R.C. Bose, "Strongly regular graphs, partial geometries and partially balanced designs", *Pacific J. Math.*, vol. 13, pp. 389–419, 1963.

[21] R.C. Bose and D.M. Mesner, "On linear associative algebras corresponding to association schemes of partially balanced designs", *Ann. Math. Statist.*, vol. 30, pp. 21–38, 1959.

[22] R.C. Bose and K.R. Nair, "Partially balanced incomplete block designs", *Sankhyā*, vol. 4, pp. 337–372, 1939.

[23] R.C. Bose and T. Shimamoto, "Classification and analysis of partially balanced incomplete block designs with two associate classes", *J. Amer. Statist. Assoc.*, vol. 47, pp. 151–184, 1952.

[24] P.G. Boyvalenkov, D.P. Danev, and S.P. Bumova, "Upper bounds on the minimum distance of spherical codes", *IEEE Trans. Inform. Theory*, vol. 42, pp. 1576–1581, 1996.

[25] P. Boyvalenkov and D. Danev, "On linear programming bounds for codes in polynomial metric spaces", to appear in *Probl. Inform. Transm.*

[26] A.E. Brouwer, A.M. Cohen, and A. Neumaier, *Distance-Regular Graphs.* Berlin: Springer-Verlag, 1989.

[27] A.E. Brouwer and T. Verhoeff, "An updated table of minimum-distance bounds for binary linear codes", *IEEE Trans. Inform. Theory*, vol. 39, pp. 662–675, 1993.

[28] A.R. Calderbank and P. Delsarte, "Extending the t-design concept", *Trans. Amer. Math. Soc.*, vol. 338, pp. 941–952, 1993.

[29] A.R. Calderbank, P. Delsarte, and N.J.A. Sloane, "A strengthening of the Assmus–Mattson theorem", *IEEE Trans. Inform. Theory*, vol. 37, pp. 1261–1268, 1991.

[30] A.R. Calderbank and J.-M. Goethals, "On a pair of dual subschemes of the Hamming scheme $H_n(q)$", *Europ. J. Combinatorics*, vol. 6, pp. 133–147, 1985.

[31] A.R. Calderbank and N.J.A. Sloane, "Inequalities for covering codes", *IEEE Trans. Inform. Theory*, vol. 34, pp. 1276–1280, 1988.

[32] P. Camion, "Codes and association schemes", in *Handbook of Coding Theory*, V.S. Pless and W.C. Huffman, Eds. Amsterdam: Elsevier, 1998.

[33] P. Camion, C. Carlet, P. Charpin, and N. Sendrier, "On correlation-immune functions", in *Advances in Cryptology–CRYPTO'91, Lecture Notes in Computer Science No. 676*, J. Feigenbaum, Ed. New York: Springer-Verlag, 1991, pp. 86-100.

[34] G.D. Cohen, M.G. Karpovsky, H.F. Mattson, Jr., and J.R. Schatz, "Covering radius—survey and recent results", *IEEE Trans. Inform. Theory*, vol. 31, pp. 328–343, 1985.

[35] P. Delsarte, "Bounds for unrestricted codes, by linear programming", *Philips Res. Repts*, vol. 27, pp. 272–289, 1972.

[36] P. Delsarte, "Four fundamental parameters of a code and their combinatorial significance", *Inform. Contr.*, vol. 23, pp. 407–438, 1973.

[37] P. Delsarte, "An algebraic approach to the association schemes of coding theory", *Philips Res. Repts Suppl.*, vol. 10, 1973.

[38] P. Delsarte, "Association schemes and t-designs in regular semilattices", *J. Combin. Theory Ser. A*, vol. 20, pp. 230–243, 1976.

[39] P. Delsarte, "Pairs of vectors in the space of an association scheme", *Philips Res. Repts*, vol. 32, pp. 373–411, 1977.

[40] P. Delsarte, "Hahn polynomials, discrete harmonics, and t-designs", *SIAM J. Appl. Math.*, vol. 34, pp. 157–166, 1978.

[41] P. Delsarte, "Bilinear forms over a finite field, with applications to coding theory", *J. Combin. Theory Ser. A*, vol. 25, pp. 226–241, 1978.

[42] P. Delsarte, "Application and generalization of the MacWilliams transform in coding theory", in *Proc. 15th Symp. Information Theory in the Benelux*, Louvain-la-Neuve, Belgium, 1994, pp. 9–44.

[43] P. Delsarte and J.-M. Goethals, "Alternating bilinear forms over $GF(q)$", *J. Combin. Theory Ser. A*, vol. 19, pp. 26–50, 1975.

[44] P. Delsarte, J.-M. Goethals, and J.J. Seidel, "Bounds for systems of lines, and Jacobi polynomials", *Philips Res. Repts,* vol. 30, pp. 91*–105*, 1975.

[45] P. Delsarte, J.-M. Goethals, and J.J. Seidel, "Spherical codes and designs", *Geom. Dedicata*, vol. 6, pp. 363–388, 1977.

[46] C.F. Dunkl, "A Krawtchouk polynomial addition theorem and wreath products of symmetric groups", *Indiana Univ. Math. J.*, vol. 25, pp. 335–358, 1976.

[47] C.F. Dunkl, "Discrete quadrature and bounds on t-design", *Mich. Math. J.*, vol. 26, pp. 81–102, 1979.

[48] C.F. Dunkl, "Orthogonal functions on some permutation groups", in *Proc. Symp. Pure Math. 34*. Providence, RI: Amer. Math. Soc., 1979, pp. 129–147.

[49] A. Erdélyi, W. Magnus, F. Oberhettinger, and F.G. Tricomi, *Higher Transcendental Functions*, vol. 2. New York: McGraw-Hill, 1953.

[50] I.A. Faradžev, A.A. Ivanov, and M.H. Klin, "Galois correspondence between permutation groups and cellular rings (association schemes)", *Graphs and Combinatorics*, vol. 6, pp. 303–332, 1990.

[51] G. Fazekas and V.I. Levenshtein, "On upper bounds for code distance and covering radius of designs in polynomial metric spaces", *J. Combin Theory Ser. A*, vol. 70, pp. 267–288, 1995.

[52] E.M. Gabidulin, "Optimal array error-correcting codes" (in Russian), *Problemy Peredachi Informatsii*, vol. 21, no. 2, pp. 103–108, 1985.

[53] R. Gangolli, "Positive definite kernels on homogeneous spaces and certain stochastic processes related to Levy's Brownian motion of several parameters", *Ann. Inst. Henri Poincaré*, vol. 3, pp. 121–226, 1967.

[54] C.D. Godsil, *Algebraic Combinatorics*. New York: Chapman & Hall, 1993.

[55] J.-M. Goethals, "Association schemes", in *Algebraic Coding Theory and Applications, CISM Courses and Lectures No. 258*, G. Longo, Ed. New York: Springer-Verlag, 1979, pp. 243–283.

[56] J.-M. Goethals and J.J. Seidel, "Orthogonal matrices with zero diagonal", *Canad. J. Math.*, vol. 19, pp. 1001–1010, 1967.

[57] A.R. Hammons, Jr., P.V. Kumar, A.R. Calderbank, N.J.A. Sloane, and P. Solé, "The $\mathbf{Z}_4$-linearity of Kerdock, Preparata, Goethals, and related codes", *IEEE Trans. Inform. Theory*, vol. 40, pp. 301–319, 1994.

[58] D.G. Higman, "Intersection matrices for finite permutation groups", *J. Algebra*, vol. 6, pp. 22–42, 1967.

[59] D.G. Higman, "Coherent configurations. Part I: Ordinary representation theory", *Geom. Dedicata*, vol. 4, pp. 1–32, 1975.

[60] S. G. Hoggar, "t-designs in projective spaces", *Europ. J. Combin.*, vol. 3, pp. 233–254, 1982.

[61] G. Hoheisel, "Über Charaktere", *Monatsch. Math. Phys.*, vol. 48, pp. 448–456, 1939.

[62] D.R. Hughes and F.C. Piper, *Design Theory*. Cambridge: Cambridge University Press, 1985.

[63] F. Jaeger, "On spin models, triply regular association schemes, and duality", *J. Algebraic Combinatorics*, vol. 4, pp. 103–144, 1995.

[64] H. Janwa, "Some new upper bounds on the covering radius of binary linear codes", *IEEE Trans. Inform. Theory*, vol. 35, pp. 110–122, 1989.

[65] G.A. Kabatyanskii and V.I. Levenshtein, "Bounds for packings on a sphere and in space", *Probl. Inform. Transm.*, vol. 14, no. 1, pp. 1–17, 1978.

[66] S. Karlin and J.L. McGregor, "The Hahn polynomials, formulas and an application", *Scripta Math.*, vol. 26, pp. 33–46, 1961.

[67] Y. Kawada, "Über den Dualitätssatz der Charaktere nichtcommutativer Gruppen", *Proc. Phys. Math. Soc. Japan (3)*, vol. 24, pp. 97–109, 1942.

[68] M. Krawtchouk, "Sur une généralisation des polynômes d'Hermite", *C. R. Acad. Sci. Paris*, vol. 189, pp. 620–622, 1929.

[69] T. Laihonen and S. Litsyn, "On upper bounds for minimum distance and covering radius of non-binary codes", *Designs, Codes and Cryptography*, vol. 14, pp. 71–80, 1998.

[70] H.W. Lenstra, Jr., "Two theorems on perfect codes", *Discr. Math.*, vol. 3, pp. 125–132, 1972.

[71] D.A. Leonard, "Orthogonal polynomials, duality and association schemes", *SIAM J. Math. Anal.*, vol. 13, pp. 656–663, 1982.

[72] D.A. Leonard, "Metric, co-metric association schemes", in *Combinatorics, Graph Theory and Computing, Proc. 15th Southeast. Conf.*, La. State Univ., Congr. Numerantium 44, 1984, pp. 277–282.

[73] V.I. Levenshtein, "On choosing polynomials to obtain bounds in packing problems" (in Russian), in *Proc. Seventh All-Union Conf. Coding Theory and Information Transmission*, Part II, Moscow, Vilnius, 1978, pp. 103–108.

[74] V.I. Levenshtein, "On bounds for packings in n-dimensional Euclidean space", *Soviet Math. Dokl.*, vol. 20, no. 2, pp. 417–421, 1979.

[75] V.I. Levenshtein, "Bounds on the maximal cardinality of a code with bounded modulus of the inner product", *Soviet Math. Dokl.*, vol. 25, no. 2, pp. 526–531, 1982.

[76] V.I. Levenshtein, "Bounds for packings of metric spaces and some of their applications" (in Russian), in *Probl. Cybern.*, vol. 40, Nauka, Moscow, 1983, pp. 43–110.

[77] V.I. Levenshtein, "Designs as maximum codes in polynomial metric spaces", *Acta Applic. Mathematicae*, vol. 29, pp. 1–82, 1992.

[78] V.I. Levenshtein, "Krawtchouk polynomials and universal bounds for codes and designs in Hamming spaces", *IEEE Trans. Inform. Theory*, vol. 41, pp. 1303–1321, 1995.

[79] V.I. Levenshtein, "Split orthogonal arrays and maximum independent resilient systems of functions", *Designs, Codes and Cryptography*, vol. 12, pp. 131–160, 1997.

[80] V.I. Levenshtein, "Equivalence of Delsarte's bounds for codes and designs in symmetric association schemes, and some applications", to appear in *Discr. Math.*

[81] V.I. Levenshtein, "Universal bounds for codes and designs", in *Handbook of Coding Theory*, V.S. Pless and W.C. Huffman, Eds. Amsterdam: Elsevier, 1998.

[82] J.H. van Lint, *Coding Theory*. Berlin: Springer-Verlag, 1982.

[83] S. Litsyn and A. Tietäväinen, "Upper bounds on the covering radius of a code with a given dual distance", *Europ. J. Combin.*, vol. 173, pp. 265–270, 1996.

[84] S.P. Lloyd, "Binary block coding", *Bell Syst. Tech. J.*, vol. 36, pp. 517–535, 1957.

[85] L. Lovász, "On the Shannon capacity of a graph", *IEEE Trans. Inform. Theory*, vol. 25, pp. 1–7, 1979.

[86] F.J. MacWilliams, "Combinatorial problems of elementary group theory", Ph.D. Thesis, Dept Mathematics, Harvard University, 1962.

[87] F.J. MacWilliams, "A theorem on the distribution of weights in a systematic code", *Bell Syst. Tech. J.*, vol. 42, pp. 79–94, 1963.

[88] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes*. New York: North-Holland, 1977.

[89] W.J. Martin and D.R Stinson, "Association schemes for ordered orthogonal arrays and (T,M,S)-nets", preprint.

[90] R.J. McEliece, E.R. Rodemich, H. Rumsey, Jr., and L.R. Welch, "New upper bounds on the rate of a code via the Delsarte–MacWilliams inequalities", *IEEE Trans. Inform. Theory*, vol. 23, pp. 157–166, 1977.

[91] R.J. McEliece, E.R. Rodemich, and H. C. Rumsey, Jr., "The Lovász bound and some generalizations", *J. Combinatorics, Information and System Sciences*, vol. 3, pp. 134–152, 1978.

[92] A.A. Nechaev, "The Kerdock code in a cyclic form" (in Russian), *Diskret. Math.*, vol. 1, no. 4, pp. 123–139, 1989. English translation in *Discr. Math. Appl.*, vol. 1, pp. 365–384, 1991.

[93] A. Neumaier, "Distances, graphs and designs", *Europ. J. Combin.*, vol. 1, pp. 163–174, 1980.

[94] A. Neumaier, "Combinatorial configurations in terms of distances", Memorandum 81-09 (Wiskunde), Eindhoven Univ. Technol., 1980.

[95] A. Neumaier, "Duality in coherent configurations", *Combinatorica*, vol. 9, pp. 59–67, 1989.

[96] A.M. Odlyzko and N.J.A. Sloane, "New bounds on the number of unit spheres that can touch a unit sphere in n dimensions", *J. Combin. Theory Ser. A*, vol. 26, pp. 210–214, 1979.

[97] R.A. Rankin, "The closest packing of spherical caps in n dimensions", *Proc. Glasgow Math. Assoc.*, vol. 2, pp. 139–144, 1955.

[98] C.R. Rao, "Factorial experiments derivable from combinatorial arrangements of arrays", *J. Roy. Statist. Soc.*, vol. 9, pp. 128–139, 1947.

[99] D.K. Ray-Chaudhuri and R.M. Wilson, "On t-designs", *Osaka J. Math.*, vol. 12, pp. 737–744, 1975.

[100] E.R. Rodemich, "An inequality in coding theory", *Amer. Math. Soc. Notices*, 773-05-39, p. 15, 1979.

[101] C. Roos, "On metric and cometric association schemes", *Delft Progr. Rep.*, vol. 4, pp. 191–220, 1979.

[102] R.M. Roth, "Maximum-rank array codes and their application to crisscross error correction", *IEEE Trans. Inform. Theory*, vol. 37, pp. 328–336, 1991.

[103] I. Schoenberg and G. Szegö, "An extremum problem for polynomials", *Composito Math.*, vol. 14, pp. 260–268, 1960.

[104] A. A. Schrijver, "A comparison of the bounds of Delsarte and Lovász", *IEEE Trans. Inform. Theory*, vol. 25, pp. 425–429, 1979.

[105] J.J. Seidel, "Strongly regular graphs", in *Surveys in Combinatorics, London Math. Soc. Lecture Notes Series No. 38*, B. Bollobàs, Ed. Cambridge: Cambridge University Press, 1979, pp. 157–180.

[106] C. Shannon, "The zero error capacity of a noisy channel", *IRE Trans. Inform. Theory*, vol. 2, pp. 8–19, 1956.

[107] C. Shannon, "Probability of error for optimal codes in Gaussian channel", *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, 1959.

[108] C. Shannon, R.G. Gallager, and E.R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels", *Inform. Contr.*, vol. 10, pp. 65–103 and 522–552, 1967.

[109] V.M. Sidelnikov, "On mutual correlation of sequences", *Soviet Math. Dokl.*, vol. 12, no. 1, pp. 197–201, 1971.

[110] V.M. Sidelnikov, "New bounds for densest packings of spheres in n-dimensional Euclidean space", *Math. USSR Sbornik*, vol. 24, pp. 147–157, 1974.

[111] V.M. Sidelnikov, "Upper bounds on the number of points of a binary code with a specified code distance", *Probl. Inform. Transm.*, vol. 10, no. 2, pp. 124–131, 1974.

[112] V.M. Sidelnikov, "Extremal polynomials used in bounds of code volume", *Probl. Inform. Transm.*, vol. 16, no. 3, pp. 174–186, 1980.

[113] T. Siegenthaler, "Correlation immunity of nonlinear combining function for cryptographic applications", *IEEE Trans. Inform. Theory*, vol. 30, pp. 776–780, 1984.

[114] L.J. Slater, *Generalized Hypergeometric Functions.* Cambridge: Cambridge University Press, 1966.

[115] N.J.A. Sloane, "An introduction to association schemes and coding theory", in *Theory and Application of Special Functions*, R.A. Askey, Ed. New York: Academic Press, 1975, pp. 225–260.

[116] N.J.A. Sloane, "Recent bounds for codes, sphere packings and related problems obtained by linear programming and other methods", *Contemp. Math.*, vol. 9, pp. 153–185, 1982.

[117] P. Solé, "A Lloyd theorem in weakly metric association schemes", *Europ. J. Combin.*, vol. 10, pp. 189–196, 1989.

[118] P. Solé and P. Stokes, "Covering radius, codimension, and dual distance width", *IEEE Trans. Inform. Theory*, vol. 39, pp. 1195–1203, 1993.

[119] D. Stanton, "Orthogonal polynomials and Chevalley groups", in *Special Functions: Group Theoretical Aspects and Applications*, R.A. Askey, Ed. Dordrecht: Reidel, 1984, pp. 87–128.

[120] D. Stanton, "An introduction to group representations and orthogonal polynomials" in *Orthogonal Polynomials*, P. Nevai, Ed., Kluwer Acad. Publ., pp. 419-433, 1990.

[121] G. Szegö, *Orthogonal Polynomials.* New York: Amer. Math. Soc., 1959.

[122] D.R. Stinson, "Combinatorial designs and cryptography" in *Surveys in Combinatorics, 1993*, K. Walter, Ed. Cambridge: Cambridge University Press, pp. 257–287, 1993.

[123] O. Tamaschke, "Zur Theorie der Permutationsgruppen mit regulärer Untergruppe I", *Math. Z.*, vol. 80, pp. 328–354, 1963.

[124] H. Tarnanen, "An approach to construct constant weight and Lee codes by using the method of association schemes", *Ann. Univ. Turku Ser. A I*, vol. 182, 1982.

[125] H. Tarnanen, M.J. Aaltonen, and J.-M. Goethals, "On the nonbinary Johnson scheme", *Europ. J. Combin.*, vol. 6, pp. 279–285, 1985.

[126] P. Terwilliger, "A characterization of P- and Q-polynomial association schemes", *J. Combin. Theory Ser. A*, vol. 45, pp. 8–26, 1987.

[127] A.A. Tietäväinen, "An upper bound on the covering radius as a function of the dual distance", *IEEE Trans. Inform. Theory*, vol. 36, pp. 1472–1474, 1990.

[128] D. Vere-Jones, "Finite bivariate distributions and semigroups of non-negative matrices", *Quart. J. Math. Oxford (2)*, vol. 22, pp. 247–270, 1971.

[129] N. Vilenkin, *Special Functions and the Theory of Group Representation.* Providence, RI: Amer. Math. Soc., 1968.

[130] H. Wang, "Two-point homogeneous spaces", *Ann. Math.*, vol. 55, pp. 177–191, 1952.

[131] L. R. Welch, R.J. McEliece, and H. Rumsey, Jr., "A low-rate improvement on the Elias bound", *IEEE Trans. Inform. Theory*, vol. 20, pp. 676–678, 1974.

[132] H. Wielandt, *Finite Permutation Groups.* New York: Academic Press, 1964.

[133] R.M. Wilson, "Inequalities for t-designs", *J. Comb. Theory Ser. A*, vol. 34, pp. 313–324, 1983.

[134] R.M. Wilson, "On the theory of t-designs", in *Enumeration and Designs*, D.M. Jackson and S.A. Vanstone, Eds. New York: Academic Press, 1984, pp. 19–49.

[135] V.A. Yudin, "Lower bounds for spherical designs", *Izvestiya: Mathematics*, vol. 61, no. 3, pp. 673–683, 1997.