Journal of Algebraic Combinatorics **12** (2000), 115–121

Uniformly Packed Codes and More Distance Regular Graphs from Crooked Functions

E.R. VAN DAM edwin.vandam@kub.nl
Department of Econometrics, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands

D. FON-DER-FLAASS* d.g.flaass@writeme.com
Institute of Mathematics, Novosibirsk 630090, Russia

Received February 23, 1999; Revised July 27, 1999

Dedicated to Jaap Seidel on his 80th birthday

Abstract. Let V and W be n-dimensional vector spaces over GF(2). A function $Q : V \to W$ is called *crooked* (a notion introduced by Bending and Fon-Der-Flaass) if it satisfies the following three properties:

$$Q(0) = 0;$$
$$Q(x) + Q(y) + Q(z) + Q(x+y+z) \neq 0 \quad \text{for any three distinct } x, y, z;$$
$$Q(x) + Q(y) + Q(z) + Q(x+a) + Q(y+a) + Q(z+a) \neq 0 \quad \text{if } a \neq 0 \ (x, y, z \text{ arbitrary}).$$

We show that crooked functions can be used to construct distance regular graphs with parameters of a Kasami distance regular graph, symmetric 5-class association schemes similar to those recently constructed by de Caen and van Dam from Kasami graphs, and uniformly packed codes with the same parameters as the double error-correcting BCH codes and Preparata codes.

Keywords: crooked function, distance-regular graph, association scheme, uniformly packed code

1. Crooked functions

Crooked functions were introduced in [1] as a means to generalise the construction of new distance regular graphs found by de Caen, Mathon, and Moorhouse [4]. In this note we show that crooked functions can similarly be used to generalise the constructions of the distance regular coset graphs of the Kasami codes (Kasami graphs) [2, Theorem 11.2.1, (13), $q = 2$], of symmetric 5-class association schemes related to Kasami graphs which were recently found by de Caen and van Dam [3], and of the double error-correcting, uniformly packed BCH (Kasami) codes and Preparata codes.

First we recall from [1] the definition and some basic properties of crooked functions, and some useful notations.

*Partly supported by the grant 96-01-01614 of the Russian Foundation for Fundamental Research. The work was partly done while the author worked at London School of Economics.

Let V and W be n-dimensional vector spaces over GF(2), and $Q : V \to W$ any mapping. We shall use the notation

$$Q(a_1, a_2, \ldots, a_m) = Q(a_1) + Q(a_2) + \cdots + Q(a_m).$$

Also, for $0 \neq a \in V$, we denote by $H_a(Q)$, or simply H_a, the set

$$H_a = H_a(Q) = \{Q(x) + Q(x + a) \mid x \in V\}.$$

We shall denote the size of a finite set X by $|X|$.

Definition 1 [1] A mapping $Q : V \to W$ is called *crooked* if it satisfies the following three properties:

(1.1) $Q(0) = 0$;

(1.2) $Q(x, y, z, x + y + z) \neq 0$ for any three distinct x, y, z;

(1.3) $Q(x, y, z, x + a, y + a, z + a) \neq 0$ if $a \neq 0$.

An equivalent but in some situations more useful description of crooked functions is given in the following proposition which was proved in [1].

Proposition 2 *If Q is a crooked mapping then*

(2.0) $n = \dim V$ *must be odd.*

(2.1) *Q is a bijection.*

(2.2) *Every set $H_a(Q)$ is the complement of a hyperplane.*

(2.3) *The sets H_a are all distinct*; *in particular*, *every complement of a hyperplane appears among them exactly once.*

Moreover, *every mapping Q satisfying property* (2.2), *and such that $Q(0) = 0$, is crooked.*

Examples of crooked functions can be constructed as follows. Let $V = W = GF(2^n)$ with n odd. Let k be a natural number coprime to n. Then the function $Q(x) = x^{1+2^k}$ is crooked. If, in the constructions to follow, we use these examples, we obtain precisely the Kasami graphs, the schemes constructed in [3], and the double error-correcting BCH codes and Preparata codes.

At present, no other examples of crooked functions are known. But the simplicity of Definition 1 suggests that many more examples should exist. Thus, an alternative title for this note might have been: "Wanted: Crooked functions. Reward increased".

2. Kasami graphs

Let $Q : V \to W$ be a crooked function; $\dim V = \dim W = n$. Let $N = 2^n = |V|$. We define a graph $K = K(Q)$ as follows. The vertex set of K is $\Omega = V \times W = \{(v, w) \mid v \in V, w \in W\}$. Vertices (v, w) and (v', w') are adjacent if and only if $v \neq v'$ and $w + w' = Q(v + v')$.

Theorem 3 *The graph $K(Q)$ is distance regular with intersection array*

$$\left\{N-1, N-2, \frac{1}{2}N+1; 1, 2, \frac{1}{2}N-1\right\}.$$

Proof: The mappings $t_{x,y}: (v, w) \to (v+x, w+y)$ for $x \in V,\ y \in W$ form a subgroup of the automorphism group of K which acts transitively on the vertices. So, it is enough to check the parameters just for one vertex, say, for $v_0 = (0, 0)$. Let, for $i = 0, 1, 2$, K_i denote the set of vertices at distance i from v_0 ($K_0 = \{v_0\}$); and let $K_3 = \Omega\backslash(K_0 \cup K_1 \cup K_2)$ (subsequently we shall see that the diameter of K is indeed 3). Also, let us denote by W_v the set $\{(v, w) \mid w \in W\}$.

It follows from the definition of K that the W_x are independent sets, and that every two distinct sets W_x, W_y are joined by a matching. Also, $K_1 = \{(x, Q(x)) \mid x \in V\backslash\{0\}\}$; $|K_1| = N-1$, as required.

Consider 2-paths from v_0 to W_a, $a \neq 0$. They all have the form

$$(0, 0) - (x, Q(x)) - (a, Q(x) + Q(x+a))$$

for $x \neq 0, a$. Thus, by Proposition (2.2), for every $h \in H_a(Q)\backslash\{Q(a)\}$ there are precisely two 2-paths from v_0 to (a, h); and this accounts for all 2-paths from v_0 to W_a. In particular, we see that K has no triangles, and that every vertex from K_2 is adjacent to precisely two vertices from K_1.

For $a \neq 0$ we have

$$K_1 \cap W_a = \{(a, Q(a))\};$$
$$K_2 \cap W_a = \{(a, h) \mid h \in H_a\backslash\{Q(a)\}\};$$
$$K_3 \cap W_a = \{(a, h) \mid h \notin H_a\}.$$

For $a = 0$, we have $W_0\backslash\{v_0\} \subset K_3$.

Let us count the number of neighbours in K_3 of an arbitrary vertex $(a, w) \in W_a$, $w \neq Q(a)$. One such neighbour can be found in W_0.

For every $b \neq 0, a$, the set $K_3 \cap W_b$ is joined by a matching to a subset $\{(a, x) \mid x \in X_b\}$ of W_a where $X_b \subset W$ is either H_b or $W\backslash H_b$: whichever of these two does not contain $Q(a)$. Let also $X_a = W\backslash H_a$; again $Q(a) \notin X_a$.

By Proposition (2.2), (2.3), the sets X_x are all possible affine hyperplanes in W not containing $Q(a)$. Therefore, every point $w \neq Q(a)$ is contained in exactly $N/2$ of them. It follows that every vertex $(a, w) \in K_2 \cap W_a$ is adjacent to $N/2+1$ vertices in K_3, and every vertex $(a, w) \in K_3 \cap W_a$ is adjacent to $(N/2-1)+1 = N/2$ vertices in K_3 (recall that $(a, w) \in K_3$ if and only if $w \in X_a$).

Neighbours of vertices from W_0 are considered similarly. Every set $K_3 \cap W_a$ for $a \neq 0$ is adjacent to the complement of a hyperplane in W_0; and each such complement occurs exactly once. Therefore each vertex in $W_0\backslash\{v_0\}$ is adjacent to $N/2$ vertices in K_3, as required.

Thus we have checked enough entries of the intersection array to conclude that it is indeed as stated in the theorem; in particular, that K is of diameter three (that is, every vertex from K_3 is adjacent to some vertex from K_2). □

3. Association schemes

For an arbitrary function $R : V \to W$ define a permutation s_R of Ω of order 2; $s_R((v, w)) = (v, w + R(v))$. The graph $s_R(K)$ is isomorphic to K; vertices (v, w) and (v', w') in it are adjacent if $w + w' = Q(v + v') + R(v) + R(v')$.

Lemma 4 *Let $R : V \to W$ be a mapping such that*

$$(\forall a \in V) \quad H_a(R) \subseteq H_a(Q).$$

Then the graphs K and $L = s_R(K)$ satisfy the following properties:

(4.1) *They are edge-disjoint.*

(4.2) *The graph $K \cup L$ has no triangles.*

(4.3) *There are no 4-tuples of vertices (x_1, x_2, x_3, x_4) such that (x_1, x_2) and (x_2, x_3) are edges of K, (x_1, x_4) and (x_4, x_3) are edges of L, and $x_1 \neq x_3$.*

Proof: The hypothesis implies that R is a bijection. Indeed, none of the sets $H_a(Q)$ contains 0; therefore $R(x) + R(y) \neq 0$ when $x \neq y$. This proves (4.1): the equalities $w + w' = Q(v + v')$ and $w + w' = Q(v + v') + R(v) + R(v')$ cannot hold simultaneously.

Suppose that vertices $x = (a, w_1), y = (b, w_2), z = (c, w_3)$ form a triangle. Neither K nor L contain triangles; so let the edges xy, yz be in K, and xz in L (the other case is similar). We have

$$Q(a + b) + Q(b + c) + Q(a + c) = R(a) + R(c).$$

This is impossible, since $Q(a+b)+Q(b+c) \in H_{a+c}(Q)$, $Q(a+c) = Q(0)+Q(a+c) \in H_{a+c}(Q)$, and $R(a) + R(c) \in H_{a+c}(R) \subseteq H_{a+c}(Q)$; but $H_{a+c}(Q)$ is sum-free.

A similar easy calculation proves (4.3). □

We shall call a mapping R satisfying the hypothesis of Lemma 4 for a crooked function Q an *accomplice* of Q. Trivially, every crooked function is an accomplice of itself.

Let R be an accomplice of a crooked function Q, and $L = s_R(K)$.

Following [3], we shall now define 5 symmetric relations $A_1, \ldots, A_5$ on Ω which will be shown to form, together with the identity relation A_0, an association scheme.

Let A_1 and A_3 be, respectively, the relations of being at distance 1 and at distance 2 in K; and A_2 and A_4, similarly, the relations of being at distance 1 and at distance 2 in L. The relation A_5 holds for vertices (v, w) and (v', w') if and only if $v = v'$ and $w \neq w'$; that is, when they lie in the same class W_v.

Theorem 5 *The relations $A_0, \ldots, A_5$ defined above form a 5-class association scheme.*

Proof: As usual, we shall identify relations with subsets of $\Omega \times \Omega$, and with their characteristic vectors viewed as $(N^2 \times N^2)$-matrices.

First let us show that $A_0 + A_1 + \cdots + A_5 = J$, the trivial relation (that is, the all-one matrix). To do this, it is sufficient to check that no pair of vertices can be in more than one of these relations—then comparing sizes does the job.

Obviously, $A_0 \cap A_i = \emptyset$ for $i \neq 0$. It is just as easy to see that $A_5 \cap A_i = \emptyset$ for $i \neq 5$. Also, we already know from the previous section that $A_1 \cap A_3 = A_2 \cap A_4 = \emptyset$. The remaining cases follow from Lemma 4: $A_1 \cap A_2 = \emptyset$ from (4.1), $A_1 \cap A_4 = A_2 \cap A_3 = \emptyset$ from (4.2), and $A_3 \cap A_4 = \emptyset$ from (4.3).

Now, following the lines of [3, Theorem 2], we prove that $A_1A_2 = A_2A_1 = A_3 + A_4 + A_5$. Consider all walks of length 2, starting at some vertex x and going first along an edge of K, and then along an edge of L. There are $(N-1)^2$ of them. By (4.1), none of them returns to x; by (4.2), none of them ends in a vertex adjacent to x; and by (4.3), they all end in distinct vertices. Since $|\Omega| = N^2 = 1 + 2(N-1) + (N-1)^2$, the claim is proved.

The relation A_5 has a very simple structure, so it is not difficult to check that, for every i, A_iA_5 is a linear combination of A_i's.

From Theorem 3 we already know that $(A_0, A_1, A_3, A_2+A_4+A_5)$ and $(A_0, A_2, A_4, A_1+A_3+A_5)$ are association schemes of distance regular graphs. Together with the equation $A_1A_2 = A_2A_1 = A_3 + A_4 + A_5$ this suffices to check that every product A_iA_j is a linear combination of A_i's with integer nonnegative coefficients (without any further resorting to the actual definition of the relations A_i). We leave this exercise to the reader. □

It is tempting to look for linear accomplices of crooked functions. Firstly, because the sets $H_a(R)$ are particularly small for a linear function: $H_a(R) = \{R(a)\}$. The second reason is that, as was mentioned in [1] just before Proposition 11, finding such a linear function would immediately give us a new example of a closed bent Kerdock set of functions, and a new Kerdock code.

Unfortunately, for known crooked functions in dimensions up to 9 there are no such linear accomplices R, as was shown by an exhaustive computer search.

4. Uniformly packed codes

Let $Q : V \to W$ be a crooked function; $\dim V = \dim W = n > 1$. Let $N = 2^n = |V|$. We define the code $C = C(Q)$ as the set of characteristic vectors of all subsets S of $V \setminus \{0\}$ such that $\sum_{r \in S} r = 0$ and $\sum_{r \in S} Q(r) = 0$. Clearly, C is a binary linear code of length $N-1$. In fact, C is a generalization of the double error-correcting BCH codes. These codes are uniformly packed, i.e., the number of codewords at distance $3\,(=e+1)$ from a word X which is at distance 2 from the code is constant, and the number of codewords at distance 3 from a word X which is at distance greater than 2 from the code is also constant.

Theorem 6 *For $n \neq 3$, the code $C(Q)$ is a double error-correcting uniformly packed code. For $n = 3$, $C(Q)$ is the perfect repetition code.*

Proof: First, suppose that there is a codeword of weight at most 4. Then there are $r_1, r_2, r_3, r_4 \in V$ such that $r_1 + r_2 + r_3 + r_4 = 0$ and $Q(r_1) + Q(r_2) + Q(r_3) + Q(r_4) = 0$. This contradicts the fact that Q is a crooked function unless all r_i are zero, so $C(Q)$ has minimum distance at least 5. Since the zero word and the all-ones word are codewords this implies that for $n = 3$, $C(Q)$ is the repetition code of length 7, and this code is perfect.

Next, consider a word X which is at distance 2 from the code $C(Q)$. We want to show that for any such X the number of codewords S at distance 3 from X is the same. Now let T be a codeword at distance 2 from X, say X and T differ in coordinates indexed by e_1 and e_2. Then $\sum_{r\in X} r = e_1 + e_2$ and $\sum_{r\in X} Q(r) = Q(e_1) + Q(e_2)$.

Suppose that S and X differ in coordinates indexed by x_1, x_2, x_3, then it becomes clear that we want to count the number of triples $\{x_1, x_2, x_3\}$ of nonzero elements of V, such that $x_1 + x_2 + x_3 = e_1 + e_2$ and $Q(x_1) + Q(x_2) + Q(x_3) = Q(e_1) + Q(e_2)$. Substituting $x_3 = x_1 + x_2 + e_1 + e_2$ in the second equation, and substituting $y = x_1 + e_1 + e_2$, $z = e_1 + e_2$, $w = Q(e_1) + Q(e_2)$, we obtain that $Q(y + z) + w = Q(x_2) + Q(x_2 + y)$. This equation has precisely two solutions for x_2 if $Q(y + z) + w \in H_y(Q)$, and otherwise it has none. Note that z and w are given. Since $Q(y + z) + Q(z) \in H_y(Q)$, we have that $Q(y+z)+w \in H_y(Q)$ if both $Q(z)$ and w are in $H_y(Q)$ or if both are not in $H_y(Q)$ (here we use that $H_y(Q)$ is the complement of a hyperplane).

Since w and $Q(z)$ are distinct and nonzero (by the properties of Q), the number of hyperplanes containing w and $Q(z)$ equals $\frac{1}{4}N - 1$, and the number of hyperplanes not containing w and $Q(z)$ equals $\frac{1}{4}N$ (this follows easily by counting). Hence by Proposition 2 it follows that the number of y such that $Q(y + z) + w \in H_y(Q)$ equals $\frac{1}{2}N - 1$, and consequently the number of triples $\{x_1, x_2, x_3\}$ with the required properties equals $\frac{1}{3}(\frac{1}{2}N - 1) - 1 = \frac{N-8}{6}$ (each triple occurs 3! times as a solution, and the solution $\{0, e_1, e_2\}$ is not allowed).

Note that the integrality of the above number of triples forces n to be odd. Also, we may now assume that $n > 3$, so that the number of triples is greater than zero, which shows that $C(Q)$ has minimum distance exactly 5.

Similarly, one can show that the number of codewords at distance 3 from a word which is at distance at least 3 from the code equals $\frac{N-2}{6}$, which completes the proof. □

Note that the proof that $C(Q)$ is a uniformly packed code goes along the same lines as the proof in [7, p. 45] that the double error-correcting BCH codes (Kasami codes) are uniformly packed. Note also that it now follows from counting that the dimension of $C(Q)$ equals $N - 1 - 2n$ (cf. [7, Thm. 1.3]).

An important consequence of the theorem is that $C(Q)$ is a double error-correcting linear code with dual degree 3 (cf. [7, Thm. 3.11]) (or is perfect in case $n = 3$), and hence it follows by the work of Delsarte (cf. [2, Chapter 11]) that the coset graph of $C(Q)$ is distance regular. Following Proposition 1 in [3] this coset graph can be reformulated as follows. Its vertex set is $V \times W$, and two distinct vertices (v, w) and (v', w') are adjacent if $w + w' = Q(v + v')$. Hence the coset graph is precisely the Kasami graph of Section 2.

Closely related to the double error-correcting BCH codes are the Preparata codes. These are binary, non-linear, double error-correcting, nearly perfect codes, that is, each word at distance at least 2 from the code has distance 2 or 3 to exactly $\frac{L}{3}$ codewords, where L is the length of the code (clearly such a code is also uniformly packed). Also here we give a generalization: by adapting the Baker-van Lint-Wilson description (cf. [6, Def. 7.4.4]) of the original Preparata code (note that other Preparata-like codes have been constructed over the ring of integers modulo four (cf. [6, Chap. 7])).

Let $P(Q)$ be the code consisting of characteristic vectors of pairs (S, T) with $S \subseteq V\backslash\{0\}$, $T \subseteq V$, such that $|T|$ is even, $\sum_{s\in S} s = \sum_{t\in T} t$, and $\sum_{s\in S} Q(s) = \sum_{t\in T} Q(t) + Q(\sum_{t\in T} t)$.

Theorem 7 *The code $P(Q)$ is a double error-correcting, nearly perfect code of size $2^{2N-2-2n}$, and length $L = 2N - 1$.*

Proof: First, note that for every choice of T, $|T|$ even, there are $|C(Q)| = 2^{N-1-2n}$ sets S such that (S, T) is a codeword (this follows by counting, and the observation that if (S, T) is a codeword, then so is $(S \div R, T)$ for every $R \in C(Q)$, where $S \div R$ stands for the symmetric difference of S and R). Thus $P(Q)$ has $2^{2N-2-2n}$ codewords.

Next, suppose that $P(Q)$ has mimimum distance at most 4, say the two codewords (S_1, T_1) and (S_2, T_2) have distance at most 4. Then it follows that S_1 and S_2 differ in 1 or 2 elements, and T_1 and T_2 differ in 2 elements (since $C(Q)$ has minimum distance 5, and T_1 and T_2 differ in an even number of elements). Without loss of generality we assume that S_1 and S_2 differ in s_1, s_2 (where we allow s_1 to be zero to cover the case where S_1 and S_2 differ in only one element), and that T_1 and T_2 differ in t_1, t_2. Now it follows that $s_1 + s_2 = t_1 + t_2$, and $Q(s_1) + Q(s_2) = Q(t_1) + Q(t_2) + Q(\sum_{t\in T_1} t) + Q(\sum_{t\in T_1} t + t_1 + t_2)$. But $Q(s_1) + Q(s_2)$, $Q(t_1) + Q(t_2)$, and $Q(\sum_{t\in T_1} t) + Q(\sum_{t\in T_1} t + t_1 + t_2) \in H_{t_1+t_2}(Q)$, which is sum-free. Hence we have a contradiction, and $P(Q)$ has minimum distance 5. It now follows from the obtained parameters that $P(Q)$ is nearly perfect (cf. [6, p. 122]). □

Added in proof. After writing this paper, we discovered the paper [5]. In this paper so-called almost bent functions are related to uniformly packed codes. In a sense, the approach in [5] is dual to ours. It follows from the results in [5] and this paper that a crooked function is almost bent. D. de Caen [private communication] showed us an easy, direct argument that this is indeed the case.

References

1. T. Bending and D. Fon-Der-Flaass, "Crooked functions, bent functions, and distance regular graphs," *Electronic Journal of Combinatorics* **5** (R34) (1998).
2. A.E. Brouwer, A.M. Cohen, and A. Neumaier, *Distance-Regular Graphs*, Springer-Verlag, Berlin, 1989.
3. D. de Caen and E.R. van Dam, "Association schemes related to Kasami codes and Kerdock sets," *Designs, Codes and Cryptography*, to appear.
4. D. de Caen, R. Mathon, and G.E. Moorhouse, "A family of antipodal distance-regular graphs related to the classical Preparata codes," *J. Alg. Combin.* **4** (1995), 317–327.
5. C. Carlet, P. Charpin, and V. Zinoviev, "Codes, bent functions and permutations suitable for DES-like cryptosystems," *Designs, Codes and Cryptography* **15** (1998), 125–156.
6. J.H. van Lint, *Introduction to Coding Theory*, 3rd edition, Springer-Verlag, 1998.
7. H.C.A. van Tilborg, "Uniformly packed codes," Thesis, Eindhoven University of Technology, 1976.