CONSTRUCTION OF LINEAR CODES HAVING PRESCRIBED PRIMAL-DUAL MINIMUM DISTANCE WITH APPLICATIONS IN CRYPTOGRAPHY

AXEL KOHNERT

Mathematical Department, University of Bayreuth, D-95440 Bayreuth, Germany
axel.kohnert@uni-bayreuth.de
http://www.mathe2.uni-bayreuth.de/people/axel.html

ABSTRACT. A method is given for the construction of linear codes with prescribed minimum distance and also prescribed minimum distance of the dual code. This works for codes over arbitrary finite fields. In the case of binary codes Matsumoto et al. showed how such codes can be used to construct cryptographic Boolean functions. This new method allows to compute new bounds on the size of such codes, extending the table of Matsumoto et al..

coding theory, minimum distance, dual minimum distance, Boolean function

1. INTRODUCTION

A linear $[n,k]_q-$code C is a $k-$dimensional subspace of the vectorspace $GF(q)^n$, where $GF(q)$ denotes the finite field with q elements. To use such a code C we work with a *generator matrix* Γ_C of C, which is a $k \times n$ matrix over $GF(q)$ whose rows are a basis of C. In coding theory we are interested in the minimum distance of the code C. For this we define the *Hamming distance* between two codewords (i.e. elements from C) $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ as the number of coordinates which are different (i.e. $u_i \neq v_i$). Then we define the *minimum distance* of C as the minimum of the Hamming distance between all pairs of codewords from C. The interest in this number comes from the fact that is possible to correct $\lfloor (d-1)/2 \rfloor$ errors if we use a code C with minimum distance d. Such a code is called an $[n,k,d]_q-$code. In this paper d will also be called primal distance.

One of the fundamental problems in coding theory [10] is the following:

Problem 1.1. For a fixed *length* n, dimension k and field $GF(q)$ find a code C with minimum distance d as large as possible.

This original problem was modified in [16] to study cryptographic problems. We denote by $C^\perp$ the dual code (i.e. the space of all words from $GF(q)^n$ which are orthogonal to all words from C) and by $d^\perp$ the minimum distance of the dual code. Now the problem in [16] can be stated as follows:

Problem 1.2. For fixed parameters n, k, q and given *primal distance* d and given *dual distance* $d^\perp$ find a code C with these properties.

Such a code is called an $[n, k, d, d^{\perp}]_q-$code. The interest in this questions comes from the fact that the generator matrix of such a code in the binary case (i.e. $q = 2$) can be used for the construction of cryptographic Boolean functions satisfying special propagation properties [14].

2. GEOMETRIC DESCRIPTION

It is known that the above problem 1.1 of finding a $[n, k]_q-$code of high minimum distance can be restated in a geometrical setting. Denote by $PG(k-1, q)$ the *finite projective geometry* of dimension $k-1$ over the finite field $GF(q)$. For our purpose we identify $PG(k-1, q)$ with the *linear lattice* of subspaces of $GF(q)^k$. The points of $PG(k-1, q)$ are the one-dimensional subspaces, the hyperplanes are the $(k-1)-$dimensional subspaces. In general an $m-$flat is the a $(m+1)-$dimensional subspace of $GF(q)^k$. The correspondence between $k-$dimensional codes over $GF(q)$ is via the columns of a generator matrix. Each column generates a one-dimensional subspace of $GF(q)^k$. This defines a correspondence ϕ between an $n-$element set of points in $PG(k-1, q)$ and an $[n, k']_q-$code where k' may be less than k. To use ϕ for our purposes we have to restrict on one side to *non-degenerate* codes (i.e. without an all-zero column in a generator matrix) and we have to allow a multiset of points in $PG(k-1, q)$ on the other side to handle the case of columns in the generator matrix, which are equal or differ only by the multiplication of a nonzero element in $GF(q)$. Then there is the well-known

Theorem 2.1. [2, 4]

There exists a non-degenerate $[n, k]_q-$ code with minimum distance at least d, if and only if there is a multiset X of size n of points in $PG(k-1, q)$ with the property:

Each hyperplane in $PG(k-1, q)$ contains at most $n-d$ points of X.

To handle the dual distance we have to use the following

Theorem 2.2. [2, 4]

Let C be a $[n, k]_q-$code C with a check matrix $\Gamma^{\perp}$. C has minimum distance greater or equal d, if and only if there are no $d-1$ columns in $\Gamma^{\perp}$ which are linearly dependent.

Then the solution of the extended problem 1.2 can be formulated using above geometric description.

Corollary 2.3.

There exists a non-degenerate $[n, k]_q-$ code with minimum distance at least d and dual distance at least $d^{\perp}$, if and only if there is a multiset X of size n of points in $PG(k-1, q)$ with the following properties:

- *each hyperplane in $PG(k-1, q)$ contains at most $n-d$ points of X.*
- *each $m-$flat contains at most $m+1$ points of X. (for all m from $0, \ldots, d^{\perp}-3$)*

The second condition is always true if we ask for dual distance 2. In this case we can get a solution which is a real multiset. In all other cases with $d^{\perp}$ greater than two, X will not be a multiset, as the second condition says for $m = 0$ that there are no multiple points. In the following we will try to construct an $[n, k, d, d^{\perp}]-$code using this geometric description.

3. DIOPHANTINE SYSTEM OF EQUATIONS

To use above characterization for the construction of codes satisfying the conditions of corollary 2.3 we restate this using a Diophantine system of equations. This was already done in [5, 6] for the case where we only prescribed the minimum distance and not also the dual distance. Denote by M^m the ($m-$flat)–point incidence matrix of $PG(k-1,q)$. The rows are labeled by the $m-$flats the columns are labeled by the points of $PG(k-1,q)$. We have

$$M^m_{i,j} = \begin{cases} 1 & \text{point } j \text{ is in flat } i \\ 0 & \text{else} \end{cases}.$$

We denote the number of rows of M^m by r_m. The number of columns is r_0. Now we can solve both problems from the introduction in Section 1 by solving a Diophantine system of equations.

Theorem 3.1. [5, 6]

There exists a non-degenerate $[n,k]_q-$ code with minimum distance at least d, if and only if there is a integral non-negative solution $x=(x_1,\ldots,x_{r_0})$ of the following Diophantine system:

- $x_1+\ldots+x_{r_0}=n$.
- $M^{k-2}x^T \leq \begin{pmatrix} n-d \\ \vdots \\ n-d \end{pmatrix}$.

where the inequality in the second part is to be read componentwise.

This Diophantine system is now enlarged by the conditions prescribing the dual distance:

Theorem 3.2.

There exists a non-degenerate $[n,k]_q-$ code with primal distance at least d and dual distance at least $d^\perp$, if and only if there is a integral non-negative solution $x=(x_1,\ldots,x_{r_0})$ of the following Diophantine system:

- $x_1+\ldots+x_{r_0}=n$.
- $M^{k-2}x^T \leq \begin{pmatrix} n-d \\ \vdots \\ n-d \end{pmatrix}$.
- $M^0 x^T \leq \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.
- $M^1 x^T \leq \begin{pmatrix} 2 \\ \vdots \\ 2 \end{pmatrix}$.
- $\vdots$
- $M^{d^\perp-3}x^T \leq \begin{pmatrix} d^\perp-2 \\ \vdots \\ d^\perp-2 \end{pmatrix}$.

4. PRESCRIBING AUTOMORPHISMS

The size of these Diophantine systems are given by the size of the corresponding projective geometry. They become too large for increasing parameters k and q to be solved directly. Like in the papers describing the solution of problem 1.1 we reduce the size of problem by prescribing automorphisms. Let G be a subgroup of $GL(k,q)$ acting on the subspaces of $GF(q)^k$. The induced action of G on the $m-$flats of $PG(k-1,q)$ gives a partition of the r_m $m-$flats into $r_{G,m}$ orbits denoted by $\omega_{G,m,1}, \omega_{G,m,2}, \ldots$. By $V_{G,m,i}$ we denote an representative of the orbit $\omega_{G,m,i}$. Then we define a condensed matrix $M^{G,m}$ by setting:

$$M^{G,m}_{i,j} := \left|\{x \in \omega_{G,0,j} : x \in V_{G,m,j}\}\right|.$$

This is a matrix with $r_{G,m}$ rows and $r_{G,0}$ columns. This matrix is well-defined as the definition is independent of the choice of the representative $V_{G,m,i}$. We get the same matrix if we add up the columns of M^m corresponding to the points in the orbit of G. The action of G is compatible with the incidence relation. This means for points p and $m-$flats V and $\phi \in G$ we have:

$$p \in V \iff \phi(p) \in \phi(V).$$

Therefore after the addition of columns the rows corresponding to $m-$flats in an orbit are equal. If we take only one copy for each orbit we get again the matrix $M^{G,m}$. This action of G on the points (= columns of a generator matrix) is used for the following definition: A linear code C has G as a group of symmetries if there is a generator matrix Γ of C whose columns correspond to full orbits of G on the 1-subspaces of $GF(q)^k$. We get a new version of theorem 3.2 using the condensed matrix:

Theorem 4.1.

There exists a non-degenerate $[n,k]_q-$ code with primal distance at least d and dual distance at least $d^\perp$and a group of symmetries which contains G as a subgroup if and only if there is a integral non-negative solution $x = (x_1, \ldots, x_{r_{G,0}})$ of the following Diophantine system:

- $\left|\omega_{G,0,1}\right| x_1 + \ldots + \left|\omega_{G,0,r_{G,0}}\right| x_{r_{G,0}} = n.$
- $M^{G,k-2}x^T \leq \begin{pmatrix} n-d \\ \vdots \\ n-d \end{pmatrix}.$
- $M^{G,0}x^T \leq \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$
- $M^{G,1}x^T \leq \begin{pmatrix} 2 \\ \vdots \\ 2 \end{pmatrix}.$
- $\vdots$
- $M^{G,d^\perp-3}x^T \leq \begin{pmatrix} d^\perp-2 \\ \vdots \\ d^\perp-2 \end{pmatrix}.$

5. RESULTS

For the binary case, which is the interesting for cryptographic applications, the authors defined in [16] the number $N(d, d^\perp)$ as the minimal length of a linear binary code with minimum distance d and dual distance $d^\perp$. They gave lower and upper bounds and computed the exact values for some combinations of the two parameters. This was done by exhaustive search. Their result was the following table:

$d\backslash d^\perp$	3	4	5	6
3	6	–		
4	7	8		
5	11	13	16	
6	12	14	17	18
7	14	15		
8	15	16		

To extend their results we first make use of the classification of small binary linear codes done by Anton Betten in [2]. Two binary codes are isomorphic if they differ only by a permutation of the coordinates. The work of Betten allows us to specify the minimum distance d and the length n, and we get (in the smaller cases) the number of different (=non-isomorphic) codes together with a generator matrix. Given such a generator matrix we compute the weight-enumerator together with the dual weight-enumerator, which we get using MacWilliams theorem. This allows us to extend the table:

$d\backslash d^\perp$	3	4	5	6	7	8
3	6	–				
4	7	8				
5	11	13	16			
6	12	14	17	18		
7	14	15	$19-20$	$20-21$	22	
8	15	16	$20-21$	$21-22$	23	24

Using the program of Ryutaroh Matsumoto for the computation of the lower bound for $N(d, d^\perp)$ given by their version of the linear programming bound in [16] we are able to show that some of the newly found codes are as short as possible. The code C found for $N(7, 7)$ is a formally self-dual code. The weight-enumerator of C and $C^\perp$ are equal. The code found for $N(8, 8)$ is a self-dual code. For larger numbers no classification results are available. But we can apply the methods described in the previous section.

Using the methods described in Section 4 we were able to construct for fixed q, n, k linear codes with prescribed distances d and $d^\perp$ for arbitrary finite fields. As an example for the non-binary case we give a table for $q = 3$ and $k = 5$ which lists for fixed $d^\perp = 4$ and all lengths n the maximum possible minimum distance d for which we were able to construct a code using our method. From the theory of caps in $PG(4, 3)$ [12] it is known, that the maximum length of code with $d^\perp = 4$ is 20.

n	6	7	8	9	10	11	12	13
d	2	2	3	4	5	6	6	6
n	14	15	16	17	18	19	20	
d	7	8	8	9	10	11	12	

Only in the case $n = 16$ this number d may not be the best possible value. There may be an other codes having primal distance 9 which we didn't found using our method. In all other cases it is known that the found minimum distance is at an upper limit, in most cases given by the Griesmer bound.

This method works for arbitrary finite fields, so we define $N_q(d, d^\perp)$ as the minimal length of a linear code over the alphabet $GF(q)$ with minimum distance d and dual distance $d^\perp$. From the constructed codes we can give upper bounds for $N_q(d, d^\perp)$. From the above table for $q = 3$ and $k = 5$ we get for example $N_3(4, 4) \leq 9$.

6. Related Work

There are several papers, which study caps [1, 4, 11, 12] in the finite projective geometry $PG(k-1, q)$. These are set of points with the additional property that on each line are at most 2 points. Now one question is which is the maximal possible size of such a point-set. If we translate the caps property into the language of the dual distance, we ask for dual distance $= 4$ but without any restrictions on the primal distance.

The reduction of the ($m-$flat)–point incidence matrix M^m using automorphisms is a general approach that works for many incidence structures for example designs [3, 15], q-analogs of designs [8], parallelisms in projective geometries [7]. The first application was in the work of Kramer and Mesner [13].

After the initial definition of the cryptographic applications it was already in the work of Carlet that he looked at the Kerdock and Preparata codes [9]. These are linear codes over the ring $\mathbb{Z}_4$. It would be interesting to apply the above method for the construction of codes with prescribed dual distance also in the case of $\mathbb{Z}_4$ and other rings.

7. Acknowledgment

The author thanks Ryutaroh Matsumoto for his helpful comments and for providing a copy of his program for the calculation of the linear programming bound from [16] which we used in section 5.

References

[1] J. Barát, Y. Edel, R. Hill, and L. Storme. On complete caps in the projective geometries over $\mathbb{F}_3$. II: New improvements. *J. Comb. Math. Comb. Comput.*, 49:9–31, 2004.

[2] Anton Betten, Michael Braun, Harald Fripertinger, Adalbert Kerber, Axel Kohnert, and Alfred Wassermann. *Error-correcting linear codes. Classification by isometry and applications. With CD-ROM.* Algorithms and Computation in Mathematics 18. Berlin: Springer. xxix, 798 p. , 2006.

[3] Anton Betten, Adalbert Kerber, Axel Kohnert, Reinhard Laue, and Alfred Wassermann. The discovery of simple 7-designs with automorphism group PΓL$(2, 32)$. Cohen, Gérard (ed.) et al., Applied algebra, algebraic algorithms and error-correcting codes. 11th international symposium, AAECC-11, Paris, France, July 17-22, 1995. Proceedings. Berlin: Springer-Verlag. Lect. Notes Comput. Sci. 948, 131-145 (1995)., 1995.

[4] Juergen Bierbrauer. *Introduction to coding theory.* Discrete Mathematics and its Applications. Boca Raton, FL: Chapman & Hall/CRC. xxiii, 390 p., 2005.

[5] M. Braun. Construction of linear codes with large minimum distance. *IEEE Transactions on Information Theory*, 50(8):1687–1691, 2004.

[6] M. Braun, A. Kohnert, and A. Wassermann. Optimal linear codes from matrix groups. *IEEE Transactions on Information Theory*, 12:4247–4251, 2005.

[7] Michael Braun. Construction of a point-cyclic resolution in PG(9,2). *Innov. Incidence Geom.*, 3:33–50, 2006.

[8] Michael Braun, Adalbert Kerber, and Reinhard Laue. Systematic construction of q-analogs of t-(v, k, λ)-designs. *Des. Codes Cryptography*, 34(1):55–70, 2005.

[9] Claude Carlet. On cryptographic propagation criteria for Boolean functions. *Inf. Comput.*, 151(1-2):32–56, 1999.

[10] Ray Hill and Emil Kolev. A survey of recent results on optimal linear codes. In *Holroyd, Fred C. (ed.) et al., Combinatorial designs and their applications. Proceedings of the one-day conference, Milton Keynes, UK, 19 March 1997. London: Chapman & Hall/CRC. Chapman & Hall/CRC Res. Notes Math. 403, 127-152* . 1999.

[11] J.W.P. Hirschfeld and L. Storme. The packing problem in statistics, coding theory and finite projective spaces: Update 2001. Blokhuis, A. (ed.) et al., Finite geometries. Proceedings of the fourth Isle of Thorns conference, Brighton, UK, April 2000. Dordrecht: Kluwer Academic Publishers. Dev. Math. 3, 201-246 (2001)., 2001.

[12] J.W.P. Hirschfeld and J.A. Thas. *General Galois geometries.* Oxford Mathematical Monographs. Oxford: Clarendon Press. xii, 407 p. , 1991.

[13] Earl S. Kramer and Dale M. Mesner. t-designs on hypergraphs. *Discrete Math.*, 15:263–296, 1976.

[14] Kaoru Kurosawa and Takashi Satoh. Design of SAC/PC(l) of order k Boolean functions and three other cryptographic criteria. *Lecture Notes in Computer Science*, 1233:434–449, 1997.

[15] Reinhard Laue, Anton Betten, and Evi Haberberger. A new smallest simple 6-design with automorphism group A_4. *Congr. Numerantium*, 150:145–153, 2001.

[16] R. Matsumoto, K. Kurosawa, T. Itoh, T. Konno, and T. Uyematsu. Primal-dual distance bounds of linear codes with application to cryptography. *IEEE Transactions on Information Theory*, 52(9):4251–4256, 2006.