Recent results on $\mathbb{Z}_4-$codes

Patrick Solé joint work with Adel Alahmadi, Tor Helleseth, Lin Sok, MinJia Shi etc

CNRS/LTCI

Paris, France, July 2016

Prologue I

In an old paper, 22 years ago, we studied Hensel lifts of BCH cyclic $\mathbb{F}_2-$codes in primitive length $2^m - 1$

In recent years we started to study cyclic $\mathbb{Z}_4-$codes in lengths $\frac{2^m-1}{N}$.

We Hensel lifted other classical binary cyclic codes like

- Melas code
- Zetterberg code
- irreducible cyclic codes

We found an algebraic decoding for the first two and constructed low correlation sequences from the last ones.

Prologue II

We also studied $\mathbb{Z}_4-$valued Boolean functions esp. bent functions.

- P. Solé, N. Tokareva, Connection between quaternary and binary bent functions `iacr.eprint.org`
- M.Shi, L. Sok, P. Solé, Classification and construction of quaternary self-dual bent functions, SETA 2016, submitted

Announcement

Inscriptions are open for CIMPA School
on
QUASI-CYCLIC and Related ALGEBRAIC CODES,
Ankara, Turkey, September 11 to 22, 2017 . Speakers include

- Buket Ozkaya : Generalized quasi-cyclic codes
- San Ling : linear quasi-cyclic codes over finite fields
- Joachim Rosenthal : convolutional codes and quasi-cyclic codes
- Roxana Smarandache : LDPC codes
- Olfa Yemen : cyclic codes leading to the notion of skew-cyclic codes

Travel grants and accomodation grants possible.

Irreducible cyclic codes

In the present work we lift the binary irreducible cyclic codes to $\mathbb{Z}_4-$ codes.
We give upper and lower bounds on the largest non-trivial correlation of their allied sequences.
As a by product we give a short proof of a claim of McEliece on the sequences of irreducible binary cyclic codes.

Irreducible cyclic binary codes

A binary cyclic code C of length n is irreducible if its parity-check polynomial is irreducible over $GF(2)$.

Trivial example : If $n = 2^m - 1$ is primitive then C is the dual of the Hamming code. Attached sequences are M-sequences.

if $n = \frac{2^m-1}{N}$, for some odd integer $N > 1$ then

Notation

Let m be an integer ≥ 2, and denote by $GR(4^m)$ the Galois ring of characteristic 4 and 4^m elements,
and by $GR(4^m)^\times$, its multiplicative group.
Let $q = 2^m$, and let N be an odd integer that divides $q-1$. We consider an irreducible cyclic $\mathbb{Z}_4-$code C_N of length $n = \frac{q-1}{N}$.
Its parity-check polynomial $H(x)$ is the minimal polynomial in $\mathbb{Z}_4[x]$ of $\beta := \gamma^N$, with γ an element of order $q-1$ of $GR(4^m)^\times$.
It can be shown that C_N is determined, up to monomial equivalence, by the multiplicative order of β in $GR(4^m)^\times$.

proof of equivalence statement

Recall that two cyclic codes are multiplier equivalent if there is an invertible element $M \in \mathbb{Z}_n$ such that such that the coordinate index permutation $x \mapsto Mx$ maps one on the other.

Claim : Codes C_N corresponding to β of the same order in $GR(4)^\times$ are multiplier equivalent.

If a, b are elements of $GR(4^m)^\times$ of the same order then there is $M \in \mathbb{Z}_n^\times$ such that $b = a^M$.

If $a = \gamma^r$, and $b = \gamma^s$ then the orders of a and b in $GR(4^m)^\times$, are, respectively, $\frac{q-1}{(q-1,r)}$ and $\frac{q-1}{(q-1,s)}$.

By hypothesis we infer that $(q-1, r) = (q-1, s) = \delta$, say. Let $r = \delta r'$ and $s = \delta s'$. Thus both r' and s' are coprime with $q-1$, hence invertible modulo n. Putting $M = \frac{r'}{s'}$ we see that M is invertible modulo n and that $r = Ms$ modulo $q-1$ hence modulo n, a divisor of $q-1$.

Trace expression

As usual the Teichmüller set T is defined by

$$T = \{0, 1, \gamma, \gamma^2, \cdots, \gamma^{q-2}\},$$

and

$$T^* = \{1, \gamma, \gamma^2, \cdots, \gamma^{q-2}\}.$$

Define, for $a \in GR(4^m)$ the evaluation vector $Ev(a)$ by the formula

$$Ev(a) = (Tr(a), Tr(a\beta), \cdots, Tr(a\beta^{n-1})).$$

The code C_N can be explicitly given as

$$C_N = \{Ev(a) |\, a \in GR(4^m)\}.$$

where Tr is the trace from $GR(4^m)$ down to $\mathbb{Z}_4$. In particular
$N = 1$ is the celebrated quaternary Kerdock code.
Its sequences were explored in 1992 by Kumar et al.

Residue code

We will denote by μ the reduction modulo 2 in $\mathbb{Z}_4$, which extends componentwise to $\mathbb{Z}_4^n$.
In particular we let $B_N = \mu(C_N)$, and observe that

$$B_N = \{ev(a) | \, a \in GF(2^m)\}.$$

where

$$ev(a) = (tr(a), tr(a\mu(\beta)), \cdots, tr(a\mu(\beta)^{n-1}))$$

and $tr(z)$ is the usual trace from $GF(2^m)$ down to $GF(2)$.
The code B_N is an irreducible binary cyclic code like in McEliece works.

Correlation

The complex correlation Θ_a attached to the generic codeword $Ev(a)$ of C_N is

$$\Theta_a = \sum_{j=0}^{n-1} i^{Tr(a\beta^j)},$$

where i is the complex fourth root of one. The (real) correlation θ_a attached to the generic codeword $ev(a)$ of B_N is

$$\theta_a = \sum_{j=0}^{n-1} (-1)^{tr(a\mu(\beta)^j)}$$

.

Sequences families

Consider a family $S = \{s_1, \cdots, s_K\}$, of K sequences with $s_i = (s_i(t))_{t=0}^{L-1}$ for $1 \leq i \leq K\}$,
each sequence of length L taking its values over $\mathbb{Z}_r$.
Let Ω be a primitive complex r^{th} root of unity .
The correlation function between the i^{th} and the j^{th} sequences is defined by $\theta_{ij}(\tau) = \sum_{t=0}^{L-1} \Omega^{s_i(t \oplus \tau) - s_j(t)}; 0 \leq \tau \leq L-1$, where $\oplus$ means addition ($\pmod{L}$).

Maximum Correlation I

Consider the maximum nontrivial correlation for complex sequences.

$$\theta_{max}(S) = \max\{|\ \theta_{ij}(\tau)\ |: 1 \leq i, j \leq K, 0 \leq \tau \leq L-1,$$

$$i \neq j \text{ if } \tau = 0\}.$$

Maximum Correlation II

We may deal with this problem also in the following way. Let $C = \{c\_i = (c_i(t))_{t=0}^{L-1} : 1 \leq i \leq KL := M\}$ be the set of the elements of S and their cyclic shifts (thus C might contain repeated elements).
Denote $\sum_{t=0}^{L-1} \omega^{c_i(t) - c_j(t)}$ by $< c\_i, c\_j >$.
Then we denote by $\theta(C)$ the quantity

$$\theta_{max}(S) = max\{| < c\_i, c\_j > | : 1 \leq i, j \leq M, i \neq j\}.$$

We say that C is an (L, M, θ) code if its length is L and its cardinality is M, and if $\theta(C)$ is less than or equal to θ.

Welch bound

The Welch bound on families of $\frac{M}{L}$ pairwise cyclically inequivalent complex sequences of period L and maximum non-trivial correlation Θ is under the form :

$$|\Theta|^2 \geq \frac{L(M-L)}{M-1}.$$

Lower bounds on the performance of the binary and quaternary sequences in this article follow then.
Open Problem : Can we use deeper bounds like Tietaivainen's or Levenshtein's ?

Lower bounds on maximum correlation

The largest non-trivial correlation attached to the code C_N is, in module, at least

$$|\theta(C_N)|^2 \geq \frac{4^m - n}{N(2^m+1)}.$$

Note that the square root of the RHS is asymptotically equivalent to $\sqrt{\frac{2^m}{N}}$ for fixed N and large m. The largest non-trivial correlation attached to the code B_N is, in module, at least

$$|\theta(B_N)|^2 \geq \frac{2^m - n}{N}.$$

Note that the square root of the RHS is asymptotically equivalent to $\frac{\sqrt{2^m(N-1)}}{N}$ for fixed N and large m.

Gauss sums over Galois rings

Define a **multiplicative** character of order N say χ by the formula $\chi(\gamma^j) = \omega^j$, where

- ω is a primitive complex root of unity of order N
- j is an integer in the range $0 \leq j \leq q-2$.

Note that χ is a character of the quotient group $\langle\gamma\rangle / \langle\beta\rangle$.
Define the Gauss sums (trivial incomplete in the sense of Langevin-Solé)

$$G_j(a) = \sum_{x \in T^*} i^{Tr(ax)} \chi^j(x),$$

for $a \in GR(4^m)$. The classical Gauss sums are then $G_j(2) = \sum_{x \in T^*} (-1)^{tr(\mu(x))} \chi^j(x)$.

Character sums and correlation

We have

$$\Theta_a = \frac{1}{N}\sum_{j=0}^{N-1} G_j(a).$$

If $a = A(1+2u)$ with $A \in T^*$ and $u \in T$ is a unit then

$$\Theta_a = \frac{1}{N}\sum_{j=0}^{N-1} G_j(1+2u)\chi^{-j}(A).$$

If a is a nonzero non unit say $a = 2\alpha$, with $\alpha \in T^*$ then

$$\Theta_a = \frac{1}{N}\sum_{j=0}^{N-1} G_j(2)\chi(\alpha)^{-j}.$$

McEliece's claim (1980)

The cyclic code B_N has an associated family of binary sequences with maximum non-trivial correlation at most $2^{\frac{m}{2}}$.
It follows from the above formula with the classical evaluation of the modulus of Gauss sums over finite fields

$$|G_j(2)| = \sqrt{q},$$

for $0 < j \leq N - 1$.
Note that $G_0(2) = -1$, by orthogonality of additive characters of $\mathbb{F}_q$.

Upper bound on the max correlation

We give an upper bound on Θ_a based on the general results of Shanbagh, Kumar, Helleseth, which are based in turn on Weil bounds for number of points of algebraic curves on finite fields. For $a \in GR(4^m)^\times$, we have

$$| 1 + N\Theta_a | \leq (2N-1)\sqrt{2^m}.$$

This comes from the bound on the character sum $\sum_{t \in T} i^{Tr(af(t))}$, with $f(t) = t^N$, that is

$$|\sum_{t \in T} i^{Tr(af(t))}| \leq (2N-1)\sqrt{2^m}.$$

Note that the preceding bound on the correlation is asymptotically equivalent to $\frac{2N-1}{N}\sqrt{2^m}$, for large m and fixed N.

Explicit expression of the correlation

With the above notation, denoting by an overbar the complex conjugation, we have

$$|G_j(1)|^2 = (2^m - 1) + (1 - i)^m \sum_{z \in T^{**}} \chi^{2j}(z) i^{Tr(\frac{z}{1+z})}.$$

Here

$$T^{**} = T^* \setminus 1 = \{\gamma, \gamma^2, \cdots, \gamma^{q-2}\}.$$

Open Problem : Compute the sum in the RHS explicitly, at least in some special cases.

Conclusion for binary sequences

In this article we have constructed based on binary irreducible cyclic codes B_N a family of binary sequences with θ_{max} in the range

$$\frac{\sqrt{2^m(N-1)}}{N} \leq \mid \theta(B_N) \mid \leq \sqrt{2^m},$$

the lower bound being asymptotic on m, for fixed N.
The period is $L = \frac{2^m-1}{N}$ and the number of cyclically non equivalent sequences is N.
We conjecture, based on our numerical data, that, for large m, the value of θ_{max} is closer to the lower than to the upper bound.

Conclusion for quaternary sequences

We also constructed based on quaternary irreducible cyclic codes a family of quaternary sequences with Θ_{max} in the range

$$\sqrt{\frac{2^m}{N}} \leq | \theta(C_N) | \leq \frac{2N-1}{N}\sqrt{2^m},$$

the bounds being asymptotic on m, for fixed N.
The period is $L = \frac{2^m-1}{N}$ and the number of cyclically inequivalent sequences is $N(2^m+1)$.
We conjecture , based on our numerical data, that, for large m, the value of θ_{max} is closer to the upper than to the lower bound.

Classical Boolean functions

A *Boolean function* in n variables is any function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$.
The set of all 2^{2^n} such functions is denoted by $\mathcal{B}_n$.
The *sign function* of f is defined as $F(x) = (-1)^{f(x)}$.
The *Walsh-Hadamard Transform* (WHT) is defined as
$W_f(u) = \sum_{x \in \mathbb{F}_2^n} (-1)^{x.u} F(x)$.
The matrix of the WHT is the Hadamard matrix H_n of Sylvester type, Let

$$H := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Let $H_n := H^{\otimes n}$ be the n-fold tensor product of H with itself. Recall the Hadamard property

$$H_n H_n^T = 2^n I_{2^n},$$

where we denote by I_M the M by M identity matrix. With these notations, $W_f(u) = H_n F$.

Classical bent functions

A function $f \in \mathcal{B}_n$, is said to be bent if $W_f(u) = \pm 2^{n/2}$ for all $u \in \mathbb{F}_2^n$.

Only exist for even n.

If f is bent, its *dual* function is defined as that element $\widehat{f}$ of $\mathcal{B}_n$ such that its sign function, henceforth denoted by $\tilde{F}$, satisfies $\tilde{F} = \frac{W_f(u)}{2^n}$.

If, furthermore, $f = \widehat{f}$, then f is self-dual bent . Similarly, if $f = \widehat{f} + 1$ then f is anti self-dual bent . Thus if f is self-dual bent, its sign function is an eigenvector of H_n associated to the eigenvalue $2^{n/2}$.

If f is anti self-dual bent, its sign function is an eigenvector of H_n associated to the eigenvalue $-2^{n/2}$.

$\mathbb{Z}_4$−bent functions

A generalized Boolean function in n variables is any function from $\mathbb{F}_2^n$ to $\mathbb{Z}_q$, for some integer q. For $q = 4$, the set of all such functions will be denoted by $\mathcal{Q}_n$.
The (complex) sign function of f is defined as $F(x) = (i)^{f(x)}$.
The quaternary Walsh-Hadamard transform $H_f(u)$ of the Boolean function f, is defined as $H_f(u) = \sum_{x\in\mathbb{F}_2^n}(-1)^{x\cdot u}F(x)$.
In matrix terms $H_f(u) = H_nF$. A function $f \in \mathcal{Q}_n$, is said to be bent if $|H_f(u)| = 2^{n/2}$ for all $u \in \mathbb{F}_2^n$. A bent quaternary function is said to be regular if there is an element $\widehat{f}$ of $\mathcal{Q}_n$, such that its sign function satisfies $H_f(u) = 2^{n/2}\tilde{F}$. If, furthermore, $f = \widehat{f}$, then f is self-dual bent. Similarly, if $f = \widehat{f} + 2$ then f is anti self-dual bent.

$\mathbb{Z}_4-$Reed-Mueller codes

There are two quaternary generalizations of Reed-Mueller codes in Hammons et al.

The codes $QRM(r, m)$ are obtained by Hensel lifting from the binary Reed-Mueller codes.

The codes $ZRM(r, m)$ are obtained by a multilevel construction from the RM codes. Symbolically,

$ZRM(r, m) = RM(r-1, m) + 2RM(r, m)$.

We require a third one, introduced in Davis and Jedwab.

Consider codes of length 2^m, generated by evaluations of quaternary Boolean functions on the 2^m points of $\mathbb{F}_2^m$. The code $RM_4(r, m)$ is generated by the monomials of order at most r. It contains $4^{\sum_{j=0}^{r}\binom{m}{j}}$ codewords and has both Hamming and Lee distance equal to 2^{m-r}

As pointed out in Borges et al. (2008),

$RM_4(r, m) = ZRM(r+1, m)$, for $r \leq m-1$.

Pairs of SD Boolean functions vs SD $\mathbb{Z}_4-$Boolean functions

Assume $F = a + bi$ is the sign function of a quaternary self-dual bent function, with a, b reals. There is a pair of binary self-dual bent functions given by their sign functions G, H as

$$\begin{aligned} G &= a + b, \\ K &= a - b. \end{aligned}$$

Conversely, every pair G, H of binary self-dual bent functions produces a quaternary self-dual bent function in that way.
$\Rightarrow$ There is no self-dual or anti-self-dual bent quaternary Boolean function in odd number of variables.

Pairs of regular bent functions vs regular $\mathbb{Z}_4$−bent function

Assume $F = a + bi$ is the sign function of a regular quaternary bent function, with a, b reals. There is a pair of binary bent functions g, k given by their sign functions G, H as

$$\begin{aligned} G &= a + b, \\ K &= a - b. \end{aligned}$$

Conversely, every pair g, k of binary bent functions produces a regular quaternary bent function in that way.
⇒There is no regular bent quaternary Boolean function in odd number of variables.

Spectral characterization

We use the notation † to denote the transconjugate of a complex valued matrix. Define the *Rayleigh quotient* attached to a complex sign function F, viewed as a column vector of length 2^n, by

$$R(F) := \frac{F^\dagger H_n F}{F^\dagger F}.$$

For $f \in \mathcal{Q}_n$, of sign function F, we have

$$-2^{n/2} \leq R(F) \leq 2^{n/2},$$

with equality in the second (resp. first) iff f is self-dual (resp. anti self-dual).

Spectral characterization : Proof

Since H_n is real symmetric, we can apply the general theory of the Rayleigh quotient of hermitian matrices.
The spectrum of H_n consists of the two eigenvalues $\pm 2^{n/2}$, with two orthogonal eigenspaces, each of dimension 2^{n-1}.
Let F^+ (resp. F^-) be the projection of F on the eigenspace attached to $2^{n/2}$ (resp. $-2^{n/2}$).
Reporting in the definition of $R(F)$ we get

$$R(F) = 2^{n/2} \frac{|F^+|^2 - |F^-|^2}{|F^+|^2 + |F^-|^2},$$

yielding the bounds

$$-2^{n/2} \leq R(F) \leq 2^{n/2},$$

where the first (resp. second) inequality is met iff $F^+ = 0$ (resp. iff $F^- = 0$).

Maiorana-McFarland type

A general class of quaternary bent functions is the following quaternary analogue of the so-called Maiorana-McFarland class. Consider all functions of the form

$$2x \cdot \phi(y) + g(y)$$

with x, y dimension $n/2$ variable vectors, ϕ any permutation in $\mathbb{F}_2^{n/2}$, and g arbitrary quaternary Boolean. In the following theorem, we consider the case where $\phi \in GL(n/2, 2)$.
A Maiorana-McFarland function is self-dual bent (resp. anti self-dual bent) if $g(y) = b \cdot y + \epsilon$ and $\phi(y) = L(y) + a$ where L is a linear automorphism satisfying $L \times L^t = I_{n/2}$, $a = L(b)$, and a has even (resp. odd) Hamming weight.
The code of parity check matrix $(I_{n/2}, L)$ is self-dual and (a, b) one of its codewords. Conversely, to the ordered pair (H, c) of a parity check matrix H of a self-dual code of length n and one of its codewords c can be attached such a Boolean function.

Dillon function type

As usual, make the convention that $\frac{1}{0} = 0$.
Assume G_0 and G_1 to be balanced Boolean function of m variables, with $G_0(0) = G_1(0) = 0$, and satisfying $\sum_{t \in \mathbb{F}_{2^m}} i^{G_0(t)+2G_1(t)} = 0$.
The quaternary Boolean function f in $2m$ variables defined by

$$f(x, y) = G_0(x/y) + 2G_1(x/y)$$

is gbent with dual

$$\widehat{f}(x, y) = G_0(y/x) + 2G_1(y/x).$$

Symmetries

In this section we derive the orbits of self-dual quaternary bent functions under the orthogonal group. Define, following Janusz, the orthogonal group of index n over $\mathbb{F}_2$ as

$$\mathcal{O}_n := \{L \in GL(n,2) \mid LL^t = I_n\}.$$

Observe that $L \in \mathcal{O}_n$ if and only if (I_n, L) is the generator matrix of a self-dual binary code of length $2n$.

The next result shows that $\mathcal{Q}_n$ is indeed wholly invariant under the group $\mathcal{O}_n$.

Let f denote a self-dual bent function in n variables.

If $L \in \mathcal{O}_n$ and $c \in \mathbb{Z}_4$ then $f(Lx) + c$ is self-dual bent.

Algorithms

Theorem Let $n \geq 2$ be an even integer and Z be arbitrary in $\{\pm 1, \pm i\}^{2^{n-1}}$. Define $Y := Z + \frac{2H_{n-1}}{2^{n/2}} Z$. If Y is in $\{\pm 1, \pm i\}^{2^{n-1}}$, then the vector (Y, Z) is the sign function of a self-dual bent function in n variables. Moreover all self-dual bent functions respect this decomposition.

There is a search algorithm for sign functions of self-dual quaternary Boolean functions, called $SDB(n, k)$ based on the above theorem, to compute all self dual quaternary bent Boolean function of degree at most k in n variables, and an analogous algorithm, called $ASDB(n, k)$ for quaternary anti-self-dual bent Boolean function in n variables, of degree at most k.

Complexity

Algorithm $SDB(n,k)$

1. Generate all $Z = i^z$ with z in $RM_4(k, n-1)$.
2. Compute all Y as $Y := Z + \frac{2H_{n-1}}{2^{n/2}} Z$.
3. If $Y \in \{\pm 1, \pm i\}^{2^{n-1}}$ output (Y, Z), else go to next Z.

To show the memory space savings with comparison with the brute force exhaustive search of complexity 2^{2^n}, the search space is only of the size of the Reed-Muller code that is $2^{2(\sum_{j=0}^{k} \binom{n-1}{j})}$.

Algorithm $ASDB(n,k)$

1. Generate all $Z = i^z$ with z in $RM_4(k, n-1)$.
2. Compute all Y as $Y := Z - \frac{2H_{n-1}}{2^{n/2}} Z$.
3. If $Y \in \{\pm 1, \pm i\}^{2^{n-1}}$ output (Y, Z), else go to next Z.

Numerics

We classify quaternary self-dual bent functions under the extended orthogonal group. Recall that two $n-$variable functions f and f' are equivalent if for any $x \in \mathbb{F}_2^n$, $f'(x) = f(Lx) + c$ for some $L \in \mathcal{O}_n, c \in \mathbb{Z}_4$.

We give the complete classification for all the functions in two and four variables,

the Gray image (the ordered pair (g, k) above) of their equivalence classes

and the classification of all quadratic functions in six variables .

In accordance with our theory, the total number of quaternary self-dual bent functions is the square of that of self-dual bent functions in Carlet et al., namely 2^2 in the case of two variables, and 20^2 in the case of four variables.