Relations between error-correcting codes and cryptographic functions

Vladimir N. Potapov

independent researcher, Siberia

(Joint work with Ferruh Özbudak)

G2A2, August 11, 2025

1. Correlation Immunity and Dual Codes
2. Sets of Functions as Codes
3. Almost Perfect Nonlinear Functions and Linear Codes

1. Correlation Immunity and Dual Codes

Functions and Codes

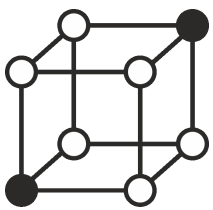
$\mathbb{F}_q$ is a finite field, where $q = p^n$, p is prime.

A function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is called a Boolean function.
A set $C \subset \mathbb{F}_2^n$ is called a binary code.
There is a one-to-one correspondence between Boolean functions and binary codes: $f^{-1}(1) = C$ or $f = \mathbf{1}_C$.

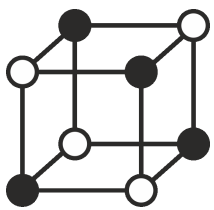
If $q > 2$ then functions $f : \mathbb{F}_q^n \to \mathbb{F}_q$ correspond to partitions $\{f^{-1}(a) : a \in \mathbb{F}_q\}$ of $\mathbb{F}_q^n$ into codes.

Correlation Immunity

An order of correlation immunity of $f : \mathbb{F}_q^n \to \mathbb{F}_q$ is equal to maximum $k = \mathrm{cor}(f)$ such that for every $a \in \mathbb{F}_q$ the cardinalities $|L \cap f^{-1}(a)|$ are the same for all $(n-k)$-dimensional face L of $\mathbb{F}_q^n$.



$\mathrm{cor}(f) = 1$

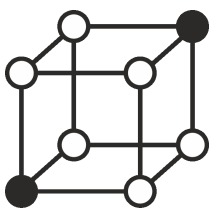


$\mathrm{cor}(f) = 2$

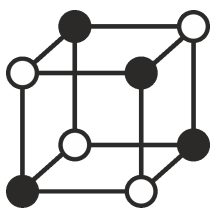
Hamming Distance

The Hamming distance $d_H(x, y)$ between $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ is the number of coordinates i where $x_i \neq y_i$. The code distance $d(C)$ of code $C \subset \mathbb{F}_q^n$ is the minimum distance between codewords $d(C) = \min_{x,y \in C} d_H(x, y)$.

Code C can correct up to $\frac{d(C)-1}{2}$ errors, and it can detect up to $d(C) - 1$ errors.



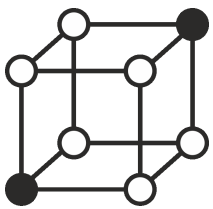
$d(C) = 3$



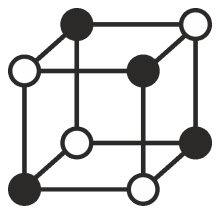
$d(C) = 2$

Dual Code

For $C \subset \mathbb{F}_q^n$ the dual code is $C^\perp = \{x \in \mathbb{F}_q^n : x_1y_1 + \cdots + x_ny_n = 0 \quad \forall y \in C\}$. If C is a linear code then $(C^\perp)^\perp = C$.



C



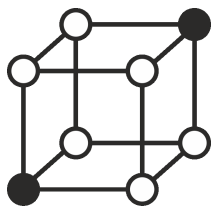
$C^\perp$

Theorem If C is a linear code then $d(C^\perp) = \mathrm{cor}(\mathbf{1}_C) + 1$ and $d(C) = \mathrm{cor}(\mathbf{1}_{C^\perp}) + 1$.

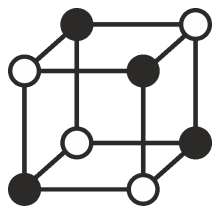
2. Sets of Functions as Codes

Set of Functions as a Code

Let $\mathcal{F} = \{f : \mathbb{F}_q^n \to \mathbb{F}_q\}$ be a set of functions. Consider f as an element of $\mathbb{F}_q^{q^n}$. Then $\mathcal{F}$ is a code in $\mathbb{F}_q^{q^n}$.



$(1, 0, 0, 0, 0, 0, 1)$



$(1, 0, 0, 1, 0, 1, 1, 0)$

Algebraic Normal Form

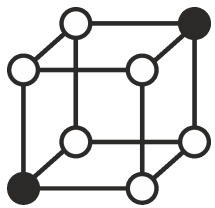
Every Boolean function can be represented in the algebraic normal form (ANF)

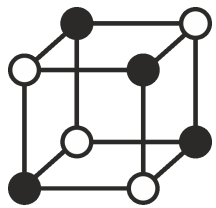
$$f(x_1, \ldots, x_n) = \bigoplus_{y \in \mathbb{F}_2^n} M_f(y) x_1^{y_1} \cdots x_n^{y_n},$$

where $x^0 = 1, x^1 = x$, $M_f : \mathbb{F}_2^n \to \mathbb{F}_2$ is the Möbius transform of f.

The weight $\mathrm{wt}(y)$ of $y \in \mathbb{F}_2^n$ is the number of nonzero coordinates of y. The algebraic degree of f is the maximum degree of the monomials in ANF, i. e., $\deg(f) = \max_{M_f(y)=1} \mathrm{wt}(y)$.

Binary Reed–Muller Code





$f_1(x_1, x_2, x_3) = 1 + x_1 + x_2 + x_3 + x_1x_2 + x_1x_3 + x_2x_3$
$f_2(x_1, x_2, x_3) = 1 + x_1 + x_2 + x_3$

$RM(r, n) = \{f : \mathbb{F}_2^n \to \mathbb{F}_2 : \deg(f) \leq r\}$.
Parameters (length, cardinality, code distance)$= (2^n, 2^k, 2^{n-r})$,
where $k = \sum_{i=0}^{r} \binom{n}{i}$ is the number of monomials.

Kerdock Code

n is even.
$\mathcal{F}(n)$ is a set consisting of n-variable quadratic nondegenerate forms f such that $f_1 + f_2$ is also nondegenerate for any $f_1, f_2 \in \mathcal{F}(n)$.
$K(n) = RM(1, n) + \mathcal{F}(n) \subset RM(2, n)$.
$\mathcal{A}_2(n) = RM(1, n)$ is the set of binary affine functions.

Parameters of $K(n)$ are $(2^n, 2^{2n}, 2^{n-1} - 2^{\frac{n}{2}-1})$.

Example: $f_1(x) = x_1x_2 + x_3x_4 + x_2x_3$, $f_2 = x_1x_3 + x_1x_4 + x_2x_4$, f_1, f_2 and $f_1 + f_2$ are nondegenerate quadratic forms.

Bent Functions

$f : \mathbb{F}_p^n \to \mathbb{F}_p$ is called a balanced function if $|f^{-1}(a)| = p^{n-1}$ for all $a \in \mathbb{F}_p$.
$f : \mathbb{F}_p^n \to \mathbb{F}_p$ is called a p-ary bent function if
$D_a f(x) = f(x + a) - f(x)$ is balanced for any $a \in \mathbb{F}_p^n \backslash \{\bar{0}\}$.
A function $f : \mathbb{F}_p^n \to \mathbb{F}_p$ is called maximum nonlinear if
$d_H(f, \mathcal{A}_p(n)) = \min\limits_{h \in \mathcal{A}_p(n)} d_H(f, h)$ is maximal.

Theorem(Rothaus(1975), V.Ryabov(2021)) For even n
1) every maximum nonlinear p-ary function is a bent function.
2) every binary bent function is maximum nonlinear.

Proposition Nondegenerate p-ary quadratic forms are bent functions if n is even.

Planar and Alltop Functions

The function $F : \mathbb{F}_q \to \mathbb{F}_q$ is called planar if for any $b \in \mathbb{F}_q \backslash \{0\}$ the derivative $D_b F(x) = F(x+b) - F(x)$ is bijective on $\mathbb{F}_q$.

A function $F : \mathbb{F}_q \to \mathbb{F}_q$ is called an Alltop function over $\mathbb{F}_q$ if, for any $a \in \mathbb{F}_q \backslash \{0\}$, the derivative $D_a F(x) = F(x+a) - F(x)$ is a planar function. Equivalently, for any $a, b \in \mathbb{F}_q \backslash \{0\}$, the expression

$$D_b D_a f(x) = f(x+a+b) - f(x+a) - f(x+b) + f(x)$$

is bijective on $\mathbb{F}_q$.

Examples: $f(x) = x^2$ is a planar function if $q \neq 2^t$;
$f(x) = x^3$ is an Alltop function if $q \neq 2^t$ and $q \neq 3^t$.

$f : \mathbb{F}_p^n \to \mathbb{F}_p$ is called p-ary Alltop function if the derivative $D_a f(x) = f(x+a) - f(x)$ is a p-ary bent for any $a \in \mathbb{F}_p^n \backslash \{\bar{0}\}$, i.e., for any $a, b \in \mathbb{F}_p^n \backslash \{\bar{0}\}$, the second-order derivative $D_b D_a f(x)$ is balanced.

Proposition Let $q = p^n$. Every coordinate function of Alltop function over $\mathbb{F}_q$ is an p-ary Alltop function.

Generalized Kerdock Codes

A set of bent functions $K = \{f : \mathbb{F}_p^n \to \mathbb{F}_p\}$ is called the Kerdock set if for any $f_1, f_2 \in K$ the difference $f_1 - f_2$ is a bent function.

Theorem (Özbudak and P. (2025+)) If f is a p-ary Alltop function then $\mathcal{K}_f = \{D_a f : a \in \mathbb{F}_p^n, \}$ is a Kerdock set.

Theorem (Özbudak and Pelen (2020)) Let $f : \mathbb{F}_p^n \to \mathbb{F}_p$ be a bent function. Then the Hamming distance between f and any affine function is not less than
$(p-1)(p^{n-1} - p^{n/2-1})$ if n is even;
$(p-1)p^{n-1} - p^{(n-1)/2}$ if n is odd.

Consider a set $\mathcal{D}_\mathcal{K} = \mathcal{K} + \mathcal{A}_p(n)$, where $\mathcal{K}$ is a Kerdock set.

Theorem (Özbudak and P. (2025+)) $\mathcal{D}_K$ is a generalized p-ary Kerdock code with parameters $(p^n, p^{2n+1}, (p-1)(p^{n-1} - p^{\frac{n}{2}-1}))$ for even n and $(p^n, p^{2n+1}, (p-1)p^{n-1} - p^{\frac{n-1}{2}})$ for odd n.

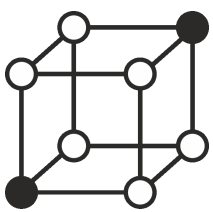
3. Almost Perfect Nonlinear Functions and Linear Codes

Parity-check Matrix

Consider a linear code $C_H = \{x \in \mathbb{F}_2^n : Hx = \bar{0}\}$ parity-check matrix H of size $k \times n$ over $\mathbb{F}_2^n$.
Rows of H belong to $C_H^{\perp}$.

Proposition If the sum of any s, $0 \leq s < S$, columns of H is not equal to zero-vector then $d(C) \geq S$.

If H consists of $2^k - 1$ different nonzero binary columns with length k then C_H is the Hamming code with parameters $(n = 2^k - 1, 2^{n-k}, 3)$.



$$H = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

Almost Perfect Nonlinear Functions

A function $F : \mathbb{F}_2^k \to \mathbb{F}_2^k$ is called an almost perfect nonlinear (APN) if for any $a \in \mathbb{F}_2^k \backslash \{\bar{0}\}$ and $b \in \mathbb{F}_2^k$ the equations $F(x) + F(x + a) = b$ have zero or two solutions $x \in \mathbb{F}_2^k$.

F is called an APN permutation if APN function F is bijective on $\mathbb{F}_2^k$.

Consider matrix M of size $2k \times (2^k - 1)$ consists of columns $\begin{pmatrix} x \\ F(x) \end{pmatrix}$, where $x \in \mathbb{F}_2^k \backslash \{\bar{0}\}$.

Theorem (Carlet, Charpin, Zinoviev (1998))
$d(C_M) = 5$ if and only if F is an APN function.

Almost Perfect Nonlinear Functions

$\mathbb{F}_2^k \sim \mathbb{F}_{2^k}$

Proposition $F(x) = x^3$ is an APN function; if k is odd then F is a permutation.

Problem Are there exist APN permutations for even $k > 6$?

Proposition APN-permutation in k variables exists if and only if there exist two Hamming codes C_1 and C_2 of length $2^k - 1$ such that $d(C_1 \cap C_2) = 5$.

Almost Perfect Nonlinear Functions

The Hamming distance between two functions is the number of arguments where values of functions are differ.

Problem Are there exist a pair of APN functions at the Hamming distance 1?

The code C of length 2^k is said a doubled Hamming code if the parity check matrix of C contains only two equal columns.

Proposition There exists a pair of APN functions at the Hamming distance 1 if and only if there exists a linear code with parameters $(n = 2^k, 2^{n-2k}, 5)$ contained a doubled Hamming code.

The End