Approximation of a combining function by functions of fewer variables

Anne Canteaut

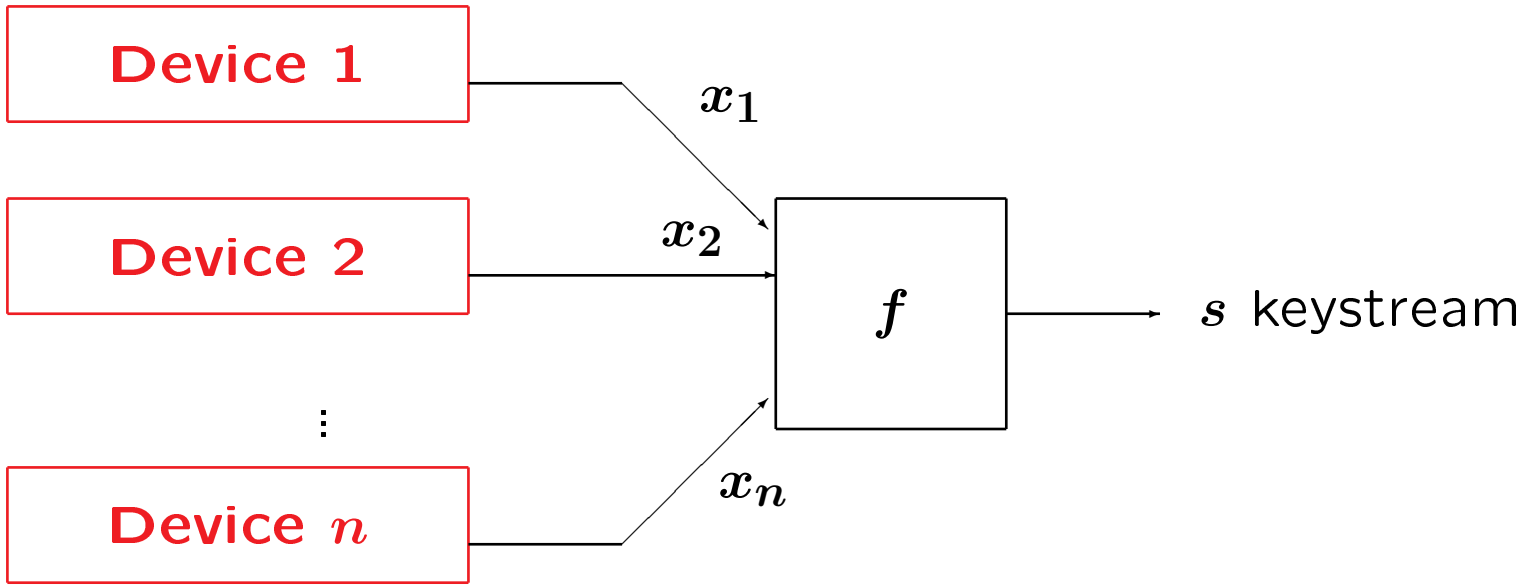
INRIA Paris-Rocquencourt
SECRET team (SEcurité, CRyptologie Et Transmissions)
Domaine de Voluceau
78153 Le Chesnay - France

ESC'08

Outline

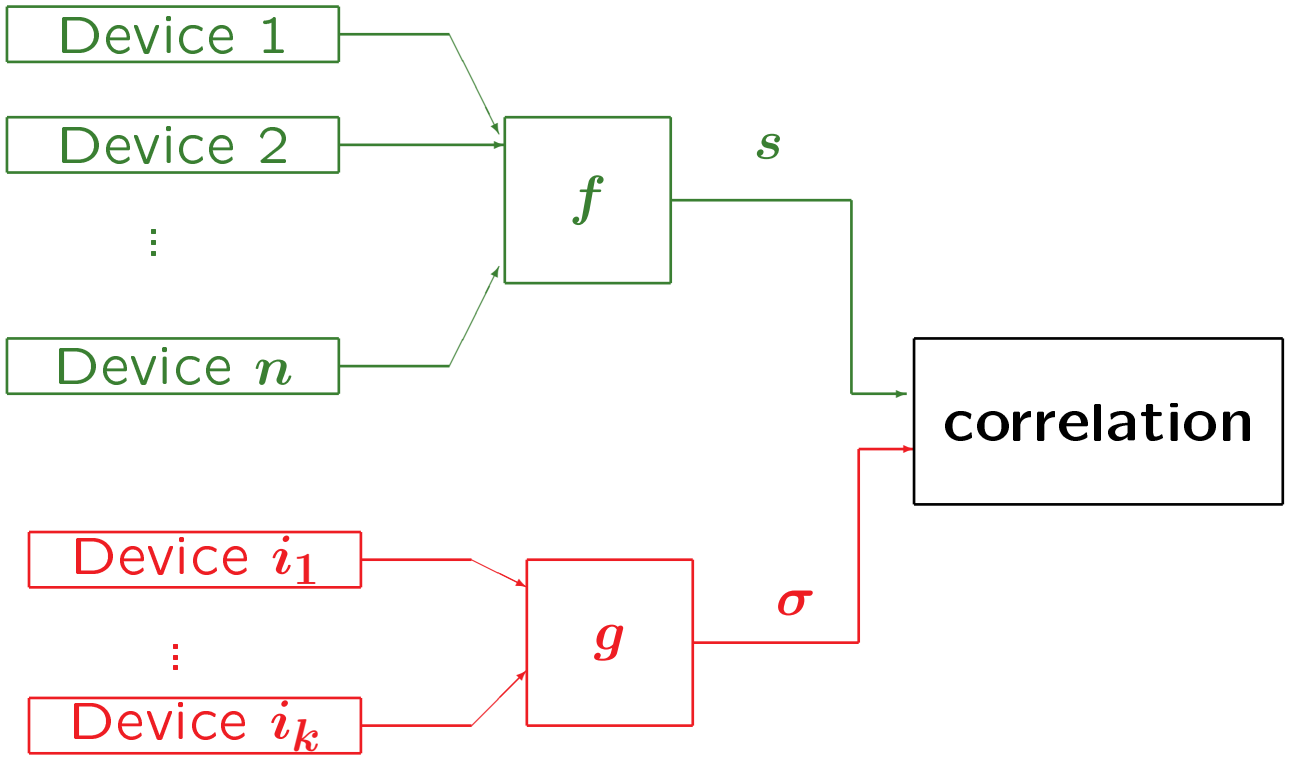
1. Divide-and-conquer attacks against some stream ciphers
2. Some attacks against Achterbahn-80
3. Link between the nonlinearity of a function and its distance to the functions of fewer variables
4. Resilient functions
5. Functions satisfying the propagation criterion
6. Bent functions

Combination generators for additive stream ciphers



where each x_i has period T_i.

Divide-and-conquer attack involving k constituent devices



where $\Pr[f(X_1, \ldots, X_n) = g(X_{i_1}, \ldots, X_{i_k})] > \frac{1}{2}$.

Resilient functions

Definition A Boolean function f is t-resilient if

$$\Pr[f(X_1, \ldots, X_n) = g(X_{i_1}, \ldots, X_{i_k})] = \frac{1}{2}$$

for any $k \leq t$ and for any function g of k variables.

The order of resiliency is the highest t such that f is t-resilient.

$\Longrightarrow$ we have to consider $t + 1$ devices together.

Building parity-check relations [Johansson-Meier-Muller 06]

Property 1. $x_1 x_2 \dots x_s$ has period $T_1 T_2 \dots T_s$.

Property 2. Let $\sigma(t) = \sum_{i=1}^{s} x_i$ and

$$\mathcal{T} = \left\{ \sum_{i=1}^{s} c_i T_i, \;\; c_i \in \{0,1\} \right\}.$$

Then, for any $t \geq 0$,

$$\sum_{\tau \in \mathcal{T}} \sigma(t+\tau) = 0.$$

Example. For $\sigma = x_1 + x_2$:

$$\sigma(t) + \sigma(t+T_1) + \sigma(t+T_2) + \sigma(t+T_1+T_2) = 0, \;\; \forall t \geq 0.$$

Building parity-check relations [Johansson-Meier-Muller 06]

Let $\sigma = g(x_{i_1}, \ldots, x_{i_k})$.

For $g = \sum_{i=1}^{m} m_i(x_{i_1}, ..., x_{i_k})$, let us consider

$$\mathcal{T} = \left\{ \sum_{i=1}^{m} c_i m_i(T_{i_1}, ..., T_{i_k}), \quad c_i \in \{0, 1\} \right\}.$$

Then,

$$\sum_{\tau \in \mathcal{T}} \sigma(t + \tau) = 0.$$

Distinguishing attack [Johansson-Meier-Muller 06]

Let $s = f(x_1, \ldots, x_n)$ where

$$\Pr[f(X_1, \ldots, X_n) = g(X_{i_1}, \ldots, X_{i_k})] = \frac{1}{2}(1 + \varepsilon) \text{ with } \varepsilon > 0.$$

For $g = \sum_{i=1}^{m} m_i(x_{i_1}, ..., x_{i_k})$ and

$$\mathcal{T} = \left\{ \sum_{i=1}^{m} c_i m_i(T_{i_1}, ..., T_{i_k}), \quad c_i \in \{0, 1\} \right\}.$$

Then,

$$\Pr\left[\sum_{\tau \in \mathcal{T}} s(t + \tau) = 0 \right] \geq \frac{1}{2}(1 + \varepsilon^{2^m}).$$

Complexity:

Time complexity $\simeq \varepsilon^{-2^{m+1}} \times 2^m$

Data complexity $\simeq \varepsilon^{-2^{m+1}} + g(T_{i_1}, \ldots, T_{i_k})$

Decimation by the period of a sequence [Hell-Johansson 06]

For $g = x_{i_j} + \sum_{i=1}^{m'} m_i(x_{i_1}, ..., x_{i_k})$, let us consider

$$\mathcal{T}' = \left\{ \sum_{i=1}^{m'} c_i m_i(T_{i_1}, ..., T_{i_k}), \quad c_i \in \{0, 1\} \right\}.$$

Then,

$$\Pr[\sum_{\tau \in \mathcal{T}'} s(t+\tau) = \sum_{\tau \in \mathcal{T}'} x_{i_j}(t+\tau)] \geq \frac{1}{2}(1 + \varepsilon^{2^{m'}}),$$

implying

$$\Pr\left[\sum_{\tau \in \mathcal{T}'} s(tT_{i_j} + \tau) = \text{ cst}\right] \geq \frac{1}{2}(1 + \varepsilon^{2^{m'}}),$$

Complexity:

Time complexity $\simeq \varepsilon^{-2^{m'+1}} \times 2^{m'}$

Data complexity $\simeq \varepsilon^{-2^{m'+1}} T_{i_j} + g'(T_{i_1}, \ldots, T_{i_k})$

Initial state recovery [Johansson-Meier-Muller 06]

For $g = \sum_{j=1}^{s} x_{i_j} + \sum_{i=1}^{m'} m_i(x_{i_1}, ..., x_{i_k})$, let us consider

$$\mathcal{T}' = \left\{ \sum_{i=1}^{m'} c_i m_i(T_{i_1}, ..., T_{i_k}), \quad c_i \in \{0,1\} \right\}.$$

Then,

$$\Pr[\sum_{\tau \in \mathcal{T}'} s(t+\tau) + \sum_{j=1}^{s} \sum_{\tau \in \mathcal{T}'} x_{i_j}(t+\tau) = 0] \geq \frac{1}{2}(1+\varepsilon^{2^{m'}}).$$

Attack:

Perform an exhaustive search for the initial states of Dev $i_1, \ldots, i_s$.
For each possible initial state, compute the parity-check equations.

Complexity:

Data complexity $\simeq \varepsilon^{-2^{m'+1}} 2 \ln 2 (L_{i_1} + \ldots + L_{i_s}) + g'(T_{i_1}, \ldots, T_{i_k})$

Time complexity $\simeq \varepsilon^{-2^{m'+1}} 2 \ln 2 (L_{i_1} + \ldots + L_{i_s}) \times 2^{m'} \times 2^{L_{i_1}+\ldots+L_{i_s}}$

Achterbahn-80 [Gammel-Göttfert-Kniffler06]

11 NLFSRs of length $L_i = 21 + i$ and of period $T_i = 2^{L_i} - 1$, $1 \leq i \leq 11$.

f: 6-resilient combining function of degree 4:

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_7 + x_9 + x_{11} + x_2x_{10} + x_2x_{11} + x_4x_8 + x_5x_6 + x_6x_8 + x_6x_{10} + x_6x_{11} + x_7x_8 + x_8x_9 + x_8x_{10} + x_9x_{10} + x_9x_{11} + x_1x_2x_8 + x_1x_4x_{10} + x_1x_4x_{11} + x_1x_8x_9 + x_1x_9x_{10} + x_1x_9x_{11} + x_2x_3x_8 + x_2x_4x_8 + x_2x_4x_{10} + x_2x_4x_{11} + x_2x_7x_8 + x_2x_8x_{10} + x_2x_8x_{11} + x_2x_9x_{10} + x_2x_9x_{11} + x_3x_4x_8 + x_3x_8x_9 + x_4x_7x_8 + x_4x_8x_9 + x_5x_6x_8 + x_5x_6x_{10} + x_5x_6x_{11} + x_6x_8x_{10} + x_6x_8x_{11} + x_7x_8x_9 + x_8x_9x_{10} + x_8x_9x_{11} + x_1x_2x_3x_8 + x_1x_2x_7x_8 + x_1x_3x_5x_8 + x_1x_3x_8x_9 + x_1x_4x_8x_{10} + x_1x_4x_8x_{11} + x_1x_5x_7x_8 + x_1x_7x_8x_9 + x_1x_8x_9x_{10} + x_1x_8x_9x_{11} + x_2x_3x_4x_8 + x_2x_3x_5x_8 + x_2x_4x_7x_8 + x_2x_4x_8x_{10} + x_2x_4x_8x_{11} + x_2x_5x_7x_8 + x_2x_8x_9x_{10} + x_2x_8x_9x_{11} + x_3x_4x_8x_9 + x_4x_7x_8x_9 + x_5x_6x_8x_{10} + x_5x_6x_8x_{11}$$

First attack against Achterbahn-80

Quadratic approximation:

$$x_1 + x_2 + x_7 + x_3x_{10} + x_4x_9, \quad \varepsilon = 2^{-5}.$$

$$\mathcal{T} = \{c_1T_3T_{10} + c_2T_4T_9, \quad c_1, c_2 \in \{0, 1\}\}$$

- Decimation by T_7
- Exhaustive search on R1 and R2.

For $\sigma = x_1 + x_2$,

$$s(tT_7) + s(tT_7 + T_3T_{10}) + s(tT_7 + T_4T_9) + s(tT_7 + T_3T_{10} + T_4T_9) =$$
$$\sigma(tT_7) + \sigma(tT_7 + T_3T_{10}) + \sigma(tT_7 + T_4T_9) + \sigma(tT_7 + T_3T_{10} + T_4T_9) + \text{cst}$$

with bias $\geq 2^{-20}$.

Data complexity $= 2^{74}$ Time complexity $= 2^{91}$.

First attack against Achterbahn-80 [Hell-Johansson06]

The exact bias of

$$s(tT_7) + s(tT_7 + T_3T_{10}) + s(tT_7 + T_4T_9) + s(tT_7 + T_3T_{10} + T_4T_9) =$$
$$\sigma(tT_7) + \sigma(tT_7 + T_3T_{10}) + \sigma(tT_7 + T_4T_9) + \sigma(tT_7 + T_3T_{10} + T_4T_9) + \mathrm{cst}$$

is not 2^{-20} but 2^{-12}.

Then,

$$\text{Data complexity } = 2^{58.3} \qquad \text{Time complexity } = 2^{75}\,.$$

Second attack against Achterbahn-80 [Naya-Plasencia06]

Linear approximation:

$$x_1 + x_2 + x_7 + (x_3 + x_{10}) + (x_4 + x_9), \quad \varepsilon = 2^{-3}.$$

$$\mathcal{T} = \{c_1 T_3 T_{10} + c_2 T_4 T_9, \quad c_1, c_2 \in \{0, 1\}\}$$

- Decimation by T_7
- Exhaustive search on R1 and R2.

For $\sigma = x_1 + x_2$,

$$s(tT_7) + s(tT_7 + T_3T_{10}) + s(tT_7 + T_4T_9) + s(tT_7 + T_3T_{10} + T_4T_9) =$$
$$\sigma(tT_7) + \sigma(tT_7 + T_3T_{10}) + \sigma(tT_7 + T_4T_9) + \sigma(tT_7 + T_3T_{10} + T_4T_9) + \text{cst}$$

with bias $\geq 2^{-12}$.

Data complexity $= 2^{58.3}$ Time complexity $= 2^{75}$.

Related issues

- Is the exact bias always given by the bias of the linear approximation?
- Can we get a better result with higher degree approximations?

General problem:

Compute
$d_H(f, \mathcal{B}_n(k))$ = distance of f to the Boolean functions depending on k variables only,

and find some properties of the best approximation.

Walsh transform of a Boolean function

Imbalance of a Boolean function:

$$\mathcal{F}(f) = \sum_{x \in \mathbf{F}_2^n} (-1)^{f(x)} = 2^n - 2wt(f) \ .$$

Linear functions: $\varphi_a \ : \ x \longmapsto a \cdot x$

Walsh (Fourier) spectrum of $f : \mathbf{F}_2^n \to \mathbf{F}_2$

$$\{\mathcal{F}(f + \varphi_a) = \sum_{x \in \mathbf{F}_2^n} (-1)^{f(x)+a \cdot x}, \ \ a \in \mathbf{F}_2^n\}$$

Nonlinearity of $f : \mathbf{F}_2^n \to \mathbf{F}_2$

Hamming distance of f to $\{\varphi_a + \varepsilon, \ a \in \mathbf{F}_2^n, \varepsilon \in \mathbf{F}_2\}$.

$$2^{n-1} - \frac{1}{2}\mathcal{L}(f) \qquad \text{where } \mathcal{L}(f) = \max_a |\mathcal{F}(f + \varphi_a)| \ .$$

Decomposition with respect to a subspace

V : subspace of $\mathbf{F}_2^n$ of dimension k

The cosets of V: $a + V$, $a \in W$ where $V \times W = \mathbf{F}_2^n$.

The decomposition of f with respect to V is the sequence $(h_1, \ldots, h_{2^{n-k}})$ of restrictions of f to $a_i + V$, $a_i \in W$,

$$h_i = f_{a_i+V}, \quad h_i : \mathbf{F}_2^k \longmapsto \mathbf{F}_2 .$$

Example. $n = 4$ and $V = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}_2^4, \quad x_1 = x_2 = 0\}$.

$x \in \mathbf{F}_2^4$	V	$(1,0,0,0) + V$	$(0,1,0,0) + V$	$(1,1,0,0) + V$
$f(x)$	0 1 0 0	0 0 1 1	0 1 0 1	0 1 1 1

Distance to $\mathcal{B}_n(K)$ by means of decompositions

Let $K \subset \{1, \ldots, n\}$. Let $V = \langle e_i, i \in K \rangle$ and $W = \langle e_i, i \notin K \rangle$.
The distance of f to the functions depending on $x_i, i \in K$ only is

$$d_H(f, \mathcal{B}_n(K)) = 2^{n-1} - \frac{1}{2} \sum_{a \in V} |\mathcal{F}(f_{a+W})| \ .$$

Moreover, the best approximation g of f is given by

$$\forall a \in V, \quad g(a) = \begin{cases} 0 & \text{if } \mathcal{F}(f_{a+W}) > 0, \\ 1 & \text{if } \mathcal{F}(f_{a+W}) < 0 \end{cases}$$

Example. $K = \{1, 2\}$ and $W = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}_2^4, \quad x_1 = x_2 = 0\}$.

$x \in \mathbf{F}_2^4$	W	$(1,0,0,0) + W$	$(0,1,0,0) + W$	$(1,1,0,0) + W$
$f(x)$	0 1 0 0	0 0 1 1	0 1 0 1	0 1 1 1
$\mathcal{F}(f_{a+W})$	2	0	0	-2
$g(a)$	0	0 or 1	0 or 1	1

$$\Longrightarrow d_H(f, \mathcal{B}_4(\{1,2\})) = 2^3 - 4/2 = 6 \ .$$

Distance to $\mathcal{B}_n(k)$ and nonlinearity

Theorem Let $K \subset \{1, \ldots, n\}$ and let $V = \langle e_i, i \in K \rangle$.

$$d_H(f, \mathcal{B}_n(K)) \geq 2^{n-1} - \frac{1}{2} \left(\sum_{\alpha \in V} \mathcal{F}^2(f + \varphi_\alpha) \right)^{\frac{1}{2}} .$$

Most notably,

$$d_H(f, \mathcal{B}_n(k)) \geq 2^{n-1} - 2^{\frac{k}{2}-1} \mathcal{L}(f)$$

where $\mathcal{L}(f) = \max_a |\mathcal{F}(f + \varphi_a)|$.

There is no accurate approximation of f by a function of a small number of variables if f has a high nonlinearity.

Approximation of a resilient function

Proposition [Xiao-Massey 88]

A Boolean function f of n variables is t-resilient if and only if

$$\forall \alpha \in \mathbf{F}_2^n,\ 0 \le w_H(\alpha) \le t,\ \ \mathcal{F}(f + \varphi_\alpha) = 0\ .$$

Theorem [C.-Trabbia 00] [Zhang 00]

Let f be t-resilient function of n variables. Then, for any K of size $t+1$,

$$d_H(f, \mathcal{B}_n(K)) = 2^{n-1} - \frac{|\mathcal{F}(f + \varphi_K)|}{2}$$

where $\varphi_K = \sum_{i \in K} x_i$, and the best approximation is achieved by the **affine function**

$$\varphi_K + \varepsilon = \sum_{i \in K} x_i + \varepsilon,\ \varepsilon \in \{0, 1\}\ .$$

Theorem Let f be t-resilient function of n variables.

$$d_H(f, \mathcal{B}_n(k)) \geq 2^{n-1} - \frac{\mathcal{L}(f)}{2} \left(\sum_{i=t+1}^{k} \binom{k}{i} \right)^{\frac{1}{2}}$$

Example. f: 6-resilient combining function of Achterbahn-80
11 variables, $\mathcal{L}(f) = 256$.

$d_H(f, \mathcal{B}_{11}(7)) = 896 \Longrightarrow \varepsilon = 2^{-3}$

$d_H(f, \mathcal{B}_{11}(8)) \geq 640 \Longrightarrow \varepsilon \leq 3 \times 2^{-3}$
(exact value: $\varepsilon = 3 \times 2^{-4}$) .

Bias of parity-checks involving $(t+1)$ variables

Theorem [C., Naya-Plasencia 07]
Let f be t-resilient function. The bias of any parity-check equation built from a $(t+1)$-variable linear approximation of f with bias ε is ε^M where M is the number of terms in the parity-check equation.

Achterbahn-128: combining function of 13 variables, 8-resilient.

Cubic approximation:

$$x_0 + x_2 + x_4 + x_9 + x_7x_{10} + x_1x_3x_{12} \text{ with bias } 2^{-6}.$$

Parity-check relation: for $\sigma = x_0 + x_2 + x_4$,
$s(tT_9)+s(tT_9+T_7T_{10})+s(tT_9+T_1T_3T_{12})+s(tT_9+T_7T_{10}+T_1T_3T_{12}) =$
$\sigma(tT_9)+\sigma(tT_9+T_7T_{10})+\sigma(tT_9+T_1T_3T_{12})+\sigma(tT_9+T_7T_{10}+T_1T_3T_{12})+$
cst.

It can be derived from

$$(x_1 + x_3 + x_{12}) + (x_7 + x_{10}) + x_9 + x_0 + x_2 + x_4 \text{ with bias } \varepsilon = 2^{-3}$$

$\Longrightarrow$ Exact bias of the parity-check relation $= \varepsilon^4 = 2^{-12}$.

Distance to $\mathcal{B}_n(n-1)$

Approximation by a function of $n-1$ variables

$$d_H(f, \mathcal{B}_n(\{1,\dots,n\} \setminus \{i\})) = 2^{n-2} - \frac{1}{4}\mathcal{F}(D_{e_i} f)$$

where $D_{e_i} f : x \mapsto f(x + e_i) + f(x)$.

Distance to $\mathcal{B}_n(n-2)$

Approximation by a function of $n-2$ variables

$$d_H(f, \mathcal{B}_n(\{1,\ldots,n\} \setminus \{i,j\})) = 2^{n-1} - 2^{n-4} - \frac{\mathcal{F}(D_{e_i} f)}{8} - \frac{\mathcal{F}(D_{e_j} f)}{8} - \frac{\mathcal{F}(D_{e_i+e_j} f)}{8} + \frac{\mathcal{F}(D_{e_i} D_{e_j} f)}{16}$$

Example. f: 6-resilient combining function of Achterbahn-80
11 variables, $\mathcal{L}(f) = 256$.

- by computing all $\mathcal{F}(D_{e_i} f)$, $1 \le i \le 11$, $d_H(f, \mathcal{B}_{11}(10)) = 128$.
- by computing all $\mathcal{F}(D_{e_i+e_j} f)$ and $\mathcal{F}(D_{e_i} D_{e_j} f)$: $d_H(f, \mathcal{B}_{11}(9)) = 704$
  $\Longrightarrow \varepsilon = 5 \times 2^{-4}$.

Approximation of a function satisfying the propagation criterion

Definition [Preneel et al 90]

f satisfies the propagation criterion of degree t ($PC(t)$) if

$$\forall a \in \mathbf{F}_2^n,\ 1 \leq w_H(a) \leq t,\ \ \mathcal{F}(D_a f) = 0\ .$$

Theorem Let f be a function of n variables satisfying $PC(t)$.

$$d_H(f, \mathcal{B}_n(k)) \geq 2^{n-1} - 2^{\frac{k}{2}-1} \left(2^n + \mathcal{M}(f) \sum_{i=t+1}^{n-k} \binom{k}{i} \right)^{\frac{1}{2}}$$

where $\mathcal{M}(f)$ is the absolute indicator of f: $\mathcal{M}(f) = \max_{e \neq 0} |\mathcal{F}(D_e f)|$.

Most notably, if $k \geq n - t$,

$$d_H(f, \mathcal{B}_n(k)) \geq 2^{n-1} - 2^{\frac{n+k}{2}-1}\ .$$

Bent functions

Let f be a function of n variables.

$$\mathcal{L}(f) = \max_{\alpha} |\mathcal{F}(f + \varphi_\alpha)| \geq 2^{\frac{n}{2}}$$

with equality if and only if f **is bent** (n even).

For any bent function f, $\mathcal{F}(f + \varphi_\alpha)$ takes 2 values only: $\pm 2^{\frac{n}{2}}$.

Dual of a bent function

The dual of f, $\widetilde{f}$, is the Boolean function of n variables defined by

$$\mathcal{F}(f + \varphi_\alpha) = 2^{n/2}(-1)^{\widetilde{f}(\alpha)} \ , \quad \alpha \in \mathbf{F}_2^n \ .$$

Approximation of a bent function

Theorem

Let f be a bent function of n variables.

Let $K \subset \{1, \ldots, n\}$ of size k.

$$d_H(f, \mathcal{B}_n(K)) \geq 2^{n-1} - 2^{\frac{n-k}{2}-1}$$

where equality holds if and only if k is even and the restriction of the dual function, $\widetilde{f}$, to $\langle e_i, i \in K \rangle$ is bent.

Conclusions

If a function f lies at high distance to the set of all affine functions, then it lies at high distance to all functions depending on a small subset of its input variables.

For a t-resilient combining function, the bias of any parity-check relation involving $(t+1)$ variables is derived from the bias of the corresponding linear approximation.