# On the computation of best second–order approximations of Boolean Functions

Nicholas Kolokotronis and Konstantinos Limniotis

Univ. of Peloponnese, Dept. of Computer Science & Tech.
Email: nkolok@uop.gr

Hellenic Data Protection Authority
Email: klimniotis@dpa.gr

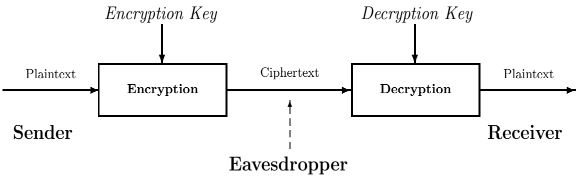
2nd Int'l Conf. on Cryptography, Network Security and Applications in the Armed Forces

April 2nd., 2014 • Hellenic Military Academy, Athens, Greece

# Talk Outline

# Symmetric ciphers

## A typical cryptosystem




## Symmetric cryptography

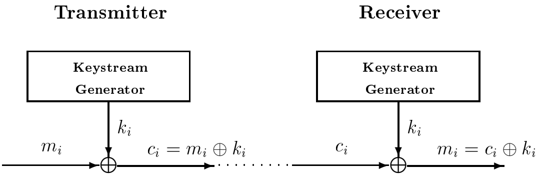
- Encryption Key = Decryption Key
- The key is only shared between the two parties

## Two types of symmetric ciphers

- Stream ciphers
- Block ciphers

# Stream ciphers

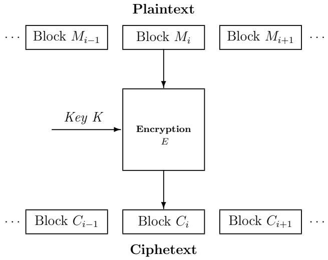
Typical Case: Binary additive stream cipher




- Suitable in environments characterized by a limited computing power or memory, and the need to encrypt at high speed
- The seed of the keystream generators constitutes the secret key
- Security depends on
  - Pseudorandomness of the keystram $k_i$
  - Properties of the underlying functions in the keystream generator

# Block ciphers

Typical Case: Electronic Codebook Mode (ECB)




- Encryption on a per-block basis (typical block size: $128$ bits)
- The encryption function $E$ performs key-dependent substitutions and permutations (Shannon's principles)
- Security depends on
  - Generation of the sub-keys used in $E$
  - Properties of the underlying functions of $E$

# A common approach for block and stream ciphers

- Despite their differences, a common study is needed for their building blocks (multi-output and single-output Boolean functions respectively)
- The attacks in block ciphers are, in general, different from the attacks in stream ciphers and vice versa. However:
  - For both cases, almost the same cryptographic criteria of functions should be in place
- Challenges:
  - There are tradeoffs between several cryptographic criteria
  - The relationships between several criteria are still unknown
  - Constructing functions satisfying all the main criteria is still an open problem

# Boolean Functions

A Boolean function $f$ on $n$ variables is a mapping from $\mathbb{F}_2^n$ onto $\mathbb{F}_2$

- The vector $f = \big(f(0,0,\ldots,0), f(1,0,\ldots,0), \ldots, f(1,1,\ldots,1)\big)$ of length $2^n$ is the truth table of $f$
- The Hamming weight of $f$ is denoted by $\mathrm{wt}(f)$
  - $f$ is balanced if and only if $\mathrm{wt}(f) = 2^{n-1}$
- The support $\mathrm{supp}(f)$ of $f$ is the set $\{\boldsymbol{b} \in \mathbb{F}_2^n : f(\boldsymbol{b}) = 1\}$

Example: Truth table of balanced $f$ with $n = 3$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x_1$ | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| $x_2$ | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| $x_3$ | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| $f(x_1, x_2, x_3)$ | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |

A vectorial Boolean function $f$ on $n$ variables is a mapping from $\mathbb{F}_2^n$ onto $\mathbb{F}_2^m$, $m > 1$

# Algebraic Normal Form and degree of functions

- Algebraic Normal Form (ANF) of $f$:

$$f(x) = \sum_{\boldsymbol{v} \in \mathbb{F}_2^n} a_{\boldsymbol{v}} x^{\boldsymbol{v}}, \quad \text{where } x^{\boldsymbol{v}} = \prod_{i=1}^{n} x_i^{v_i}$$

.

  - The sum is performed over $\mathbb{F}_2$ (XOR addition)
- The degree $\deg(f)$ of $f$ is the highest number of variables that appear in a product term in its ANF.
- If $\deg(f) = 1$, then $f$ is called affine function
  - If, in addition, the constant term is zero, then the function is called linear
- In the previous example: $f(x_1, x_2, x_3) = x_1 x_2 + x_2 x_3 + x_1$.
- $\deg(f) = 2$

# Univariate representation of Boolean functions

- $\mathbb{F}_2^n$ is isomorphic to the finite field $\mathbb{F}_{2^n}$,
- $\Rightarrow$ Any function $f \in \mathbb{B}_n$ can also be represented by a univariate polynomial, mapping $\mathbb{F}_{2^n}$ onto $\mathbb{F}_2$, as follows

$$f(x) = \sum_{i=0}^{2^n-1} \beta_i x^i$$

where $\beta_0, \beta_{2^n-1} \in \mathbb{F}_2$ and $\beta_{2i} = \beta_i^2 \in \mathbb{F}_{2^n}$ for $1 \leq i \leq 2^n - 2$

- The coefficients of the polynomial determine the Discrete Fourier Transform of $f$
- The degree of $f$ can be directly deduced by the univariate representation
- The univariate representation is more convenient in several cases

# Walsh transform

## Definition

The Walsh transform $\widehat{\chi}_f(a)$ at $a \in \mathbb{F}_2^n$ of $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is

$$\widehat{\chi}_f(a) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+ax^T} = 2^n - 2\,\mathrm{wt}(f + \phi_a)$$

where $\phi_a(x) = ax^T = a_1x_1 + \cdots + a_nx_n$

- Computational complexity: $\mathcal{O}(n2^n)$ (via fast Walsh transform)
- Parseval's theorem: $\sum_{a \in \mathbb{F}_2^n} \widehat{\chi}_f(a)^2 = 2^{2n}$

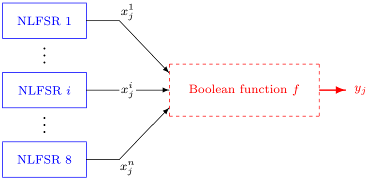
# Linear approximation attacks

- Cryptographic functions need to be balanced, as well as of high degree
  - The maximum possible degree of a balanced Boolean function with $n$ variables is $n-1$
- High degree though is not adequate to prevent linear cryptanalysis (in block ciphers - Matsui, 1992) or best affine approximation attacks (in stream ciphers - Ding et. al., 1991)
  - A function should not be well approximated by a linear/affine function
  - Any function of degree $1$ that best approximates $f$ is a best affine/linear approximation of $f$

# Example of approximation attacks

The Achterbahn cipher [Gammel-Göttfert-Kniffler,2005] (candidate in eSTREAM project)

- Stream cipher, based on a nonlinear combination generator




- Lengths of nonlinear FSRs: 22-31
- $f(x_1, \ldots, x_8) = \sum_{i=1}^{4} x_i + x_5x_7 + x_6x_7 + x_6x_8 + x_5x_6x_7 + x_6x_7x_8$
- Johansson-Meier-Muller, 2006: cryptanalysis via the linear approximation $g(x_1, \ldots, x_8) = x_1 + x_2 + x_3 + x_4 + x_6$, satisfying wt$(f + g) = 64$ $(p(f = g) = 3/4)$

# The notion of nonlinearity

- The minimum distance between $f$ and all affine functions is the nonlinearity of $f$:

$$\mathsf{nl}(f) = \min_{l \in \mathbb{B}_n : \deg(l) = 1} \mathsf{wt}(f + l)$$

- Relathionship with Walsh transform

$$\mathsf{nl}(f) = 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_2^n} |\widehat{\chi}_f(a)|$$

  - $\Rightarrow$ Nonlinearity is computed via the Fast Walsh Transform
- High nonlinearity is prerequisite for thwarting attacks based on affine (linear) approximations

# Known results on nonlinearity of Boolean functions

- For even $n$, the maximum possible nonlinearity is $2^{n-1} - 2^{n/2-1}$, achieved by the so-called bent functions
  - Many constructions are known (not fully classified yet)
  - But bent functions are never balanced!
- For odd $n$, the maximum possible nonlinearity is still unknown
  - By concatenating bent functions, we can get nonlinearity $2^{n-1} - 2^{\frac{n-1}{2}}$. Can we impove this?
    - For $n \leq 7$, the answer is no
    - For $n \geq 15$, the answer is yes (Patterson-Wiedemann, 1983 - Dobbertin, 1995 - Maitra-Sarkar, 2002)
    - For $n = 9, 11, 13$, such functions have been found more recently (Kavut, 2006)
- Several constructions of balanced functions with high nonlinearity exist (e.g. Dobbertin, 1995). However:
  - Finding the highest possible nonlinearity of balanced Boolean functions is still an open problem

# The Maiorana-McFarland class of functions

- A widely known class of functions with nice cryptographic properties
- $f \in \mathbb{B}_{k+s}$ satisfying the following:

$$f(y, x) = F(y)x + h(y),\ x \in \mathbb{F}_2^k,\ y \in \mathbb{F}_2^s$$

  - $F$ is any mapping from $\mathbb{F}_2^k$ to $\mathbb{F}_2^s$
  - $h \in \mathbb{B}_s$
- If $k = s$ and $F$ is a permutation over $\mathbb{F}_2^k \Rightarrow f$ is bent (e.g. Dillon, 1974)
- For injective $F$, if $\mathrm{wt}(F(\tau)) \geq t + 1$ for all $\tau \in \mathbb{F}_2^s$, then $f$ is $t$-resilient - i.e. resistant against correlation attacks (Camion et. al., 1992).

# Higher-order nonlinearity

- Approximating a function by a low-order function (not necessarily linear) may also lead to cryptanalysis (Non–linear cryptanalysis - Knudsen-1996, low-order approximation attacks - Kurosawa et. al. - 2002)
- The $r$th order nonlinearity of a Boolean function $f \in \mathbb{B}_n$ is given by

$$\mathsf{nl}_r(f) = \min_{g \in \mathbb{B}_n : \deg(g) \leq r} \mathsf{wt}(f+g)$$

- The $r$th order nonlinearity remains unknown for $r > 1$
  - Recursive lower bounds on $\mathsf{nl}_r(f)$ (Carlet, 2008)
  - Specific lower and upper bounds for $\mathsf{nl}_2(f)$ (Cohen, 1992 - Carlet, 2007)
  - More recent lower bounds for $2$-nd order nonlinearity: Gangopadhyay et. al. - 2010, Garg et. al. - 2011, Singh - 2011, Singh et. al. - 2013

# Problem Statement

## What has been done?

- $r$th order nonlinearity remains unknown, for $r \geq 2$
- No much is known regarding constructions of functions with high $r$-th nonlinearity, for $r \geq 2$
- Even if $r$-th order nonlinearity is estimated, finding best $r$-th order approximations is a difficult task (even for $r = 2$)

## How do we proceed?

- Link internal structure with $r$th order nonlinearity
- Examine cubic functions of specific form
  - Use of properties of the underlying quadratic functions
- Use of perfect nonlinear mappings to achieve high second–order nonlinearity

# Computing best 2nd–order approximations

- Efficient solution for specific class of 3-rd degree functions (Kolokotronis-Limniotis-Kalouptsidis, 2007)
  - The problem is appropriately reduced in computing best affine approximation attacks of the underlying 2-nd degree sub-functions
- The simplest case: There is a common variable $x_i$ in all cubic terms of $f \in \mathbb{B}_n$
- Decompose $f$ into quadratic $f_0, f_1 \in \mathbb{B}_{n-1}$: $f = (1 + x_i)f_0 + x_i f_1$
- Fixing either $f_0$ or $f_1$, and appropriately modifying the other, gives a best 2nd–order approximation
- The problem of computing best 2nd–order approximations of cubic functions is reduced to finding best linear approximations of quadratic functions (which is an easy task)

# A simple example

$$
\begin{tikzcd}[column sep=small]
& f(x_1,\ldots,x_5) = x_1x_2x_4 + x_2x_3x_5 + x_1x_5 + x_2x_3 + x_3x_4 + x_3x_5 + x_3 \arrow[dl, "1+x_2"'] \arrow[dr, "x_2"] & \\
f_0 = \underbrace{x_1x_5 + x_3x_4 + x_3x_5}_{q^0} + x_3 & & f_1 = \underbrace{x_1x_4 + x_1x_5 + x_3x_4}_{q^1}
\end{tikzcd}
$$

**Case 1**: Fix $f_0$

$$
\begin{tikzcd}
& \xi_f^0 \arrow[dl, "1+x_2"'] \arrow[dr, "x_2"] & \\
f_0 & & q^0 + \lambda_{\underbrace{x_1x_4 + x_3x_5}_{q^0+q^1}}
\end{tikzcd}
$$

**Case 2**: Fix $f_1$

$$
\begin{tikzcd}
& \xi_f^1 \arrow[dl, "1+x_2"'] \arrow[dr, "x_2"] & \\
q^1 + x_3 + \lambda_{\underbrace{x_1x_4 + x_3x_5}_{q^0+q^1}} & & f_1
\end{tikzcd}
$$

- A best linear approximaton of $q_2 = q_0 + q_1 = x_1x_4 + x_3x_5$ is the all-zeroes function. Then:
  - $\xi_f^0 = x_1x_5 + x_3x_4 + x_3x_5 + x_2x_3 + x_3$
  - $\xi_f^1 = x_1x_4 + x_1x_5 + x_3x_4 + x_2x_3 + x_3$
- $\mathrm{nl}_2(f) = \mathrm{nl}(q_0 + q_1)$, where $f \in \mathbb{B}_n$, $q_0 + q_1 \in \mathbb{B}_{n-1}$

# Practical application

- Recall Achterbahn's combiner function:

$$f(x_1,\ldots,x_8)=\sum_{i=1}^{4}x_i+x_5x_7+x_6x_7+x_6x_8+x_5x_6x_7+x_6x_7x_8$$

- $x_6$ is common in all cubic terms
- $q(x)=x_5x_7+x_6x_8+x_1+x_2+x_3+x_4$ is a best 2-nd approximation
  - Efficiently computed via the aforementioned procedure
  - All others best approximations can also be computed
- $\mathrm{wt}(f+q)=32\ (p(f=q)=7/8>3/4)$

# Generalization of the results

- Generalization to separable $3$-rd degree functions (Kolokotronis-Limniotis-Kalouptsidis, 2009)
- $f = f_1 + \cdots + f_m$ where $f_1, \ldots, f_m$ are defined cubic terms defined on disjoint sets of variables.
- All the best $2$nd–order approximations are efficiently computed
- Large values of $m$ increase $2$nd–order nonlinearity
  - Seaparability though seems to pose a risk from a cryptographic point of view
  - The first class of functions whose best $2$nd–order approximations can be efficiently found

# A case of highly nonlinear function $f$ with $\mathsf{nl}_2(f) = \mathsf{nl}(f)$

## Cubic functions in the general Maiorana–McFarland class

$$f(x, y) = F(x)y^{T}, \qquad (x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^m .$$

- $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$: quadratic vectorial Perfect Nonlinear (PN) function
  - All linear combinations of the $m$ underlying Boolean functions are bent
- (Kolokotronis-Limniotis, 2012): Let $f \in \mathbb{B}_{n+m}$ be cubic function of the above form, where each linear combination of the Boolean functions of $F$ is bent of minimal weight, and $m \leq \lfloor \frac{n}{4} \rfloor$. Then,

$$\mathsf{nl}_2(f) = 2^{n+m-1} - 2^{n/2-1}(2^{n/2} + 2^m - 1) = \mathsf{nl}(f)$$

- Best $2$-nd order approximations are also efficiently computed
  - Each linear combination of the output columns of $F$

# Bounds on the Second Order Nonlinearity

| $n$ | KL12 | C08 | GST10 | GG11 | GG09 | LHG10 | S11 | SW09 | SW11 |
|---|---|---|---|---|---|---|---|---|---|
| 5 | 6 | 6 | – | 4 | 5 | 6 | 1 | 4 | 4 |
| 6 | 12 | 12 | 15 | 10 | 10 | 16 | 10 | 17 | 8 |
| 7 | 28 | 36 | 30 | 20 | 32 | 36 | 19 | 34 | 16 |
| 8 | 56 | 72 | 60 | 52 | 64 | 78 | 64 | 84 | 62 |
| 9 | 120 | 176 | 120 | 104 | 166 | 166 | 128 | 168 | 124 |
| 10 | 360 | 352 | 378 | 256 | 331 | 351 | 330 | 386 | 248 |
| 11 | 720 | 802 | 756 | 512 | 768 | 737 | 661 | 772 | 496 |
| 12 | 1488 | 1604 | 1524 | 1187 | 1536 | 1536 | 1535 | 1689 | 1318 |
| 13 | 2976 | 3468 | 3048 | 2374 | 3372 | 3184 | 3071 | 3378 | 2636 |
| 14 | 6048 | 6936 | 7139 | 5296 | 6744 | 6567 | 6742 | 7172 | 5272 |
| 15 | 14112 | 14605 | 14278 | 10592 | 14336 | 13488 | 13485 | 14344 | 10544 |
| 16 | 28224 | 29210 | 28556 | 23027 | 28672 | 27608 | 28669 | 29877 | 24561 |
| 17 | 56896 | 60517 | 57112 | 46054 | 59744 | 56341 | 57341 | 59754 | 49122 |
| 18 | 113792 | 121034 | 122758 | 98304 | 119487 | 114688 | 119482 | 122888 | 98244 |
| 19 | 228480 | 247951 | 245516 | 196608 | 245760 | 232952 | 238968 | 245776 | 196488 |
| 20 | 489600 | 495902 | 491278 | 414071 | 491520 | 472273 | 491513 | 501129 | 431562 |

# Summary

## Significance of our results

- The second–order nonlinearity of the Maiorana–McFarland class outer–performs the second–order nonlinearity of other known constructions.
  - This class is further strengthened in terms of cryptographic properties
- Best quadratic approximations can be efficiently computed.
  - Further extension of the results of [KLK-09] - non-separable cases

## Concluding remarks

- Constructions based on perfect nonlinear mappings seem to be the right way to obtain functions with high first–order and second–order nonlinearity.

# Further research

## Further Research

- Examine the functions considered so far via the univariate representation
  - This representation seems, in many other cases, to be more convenient
  - How the separability property is being reflected into the univariate representation?
  - Extension of the results achieved so far
- Study trade offs between $r$–th order nonlinearity and other cryptographic criteria

# Questions & Answers

Thank you for your attention!