Counting first order correlation-immune functions

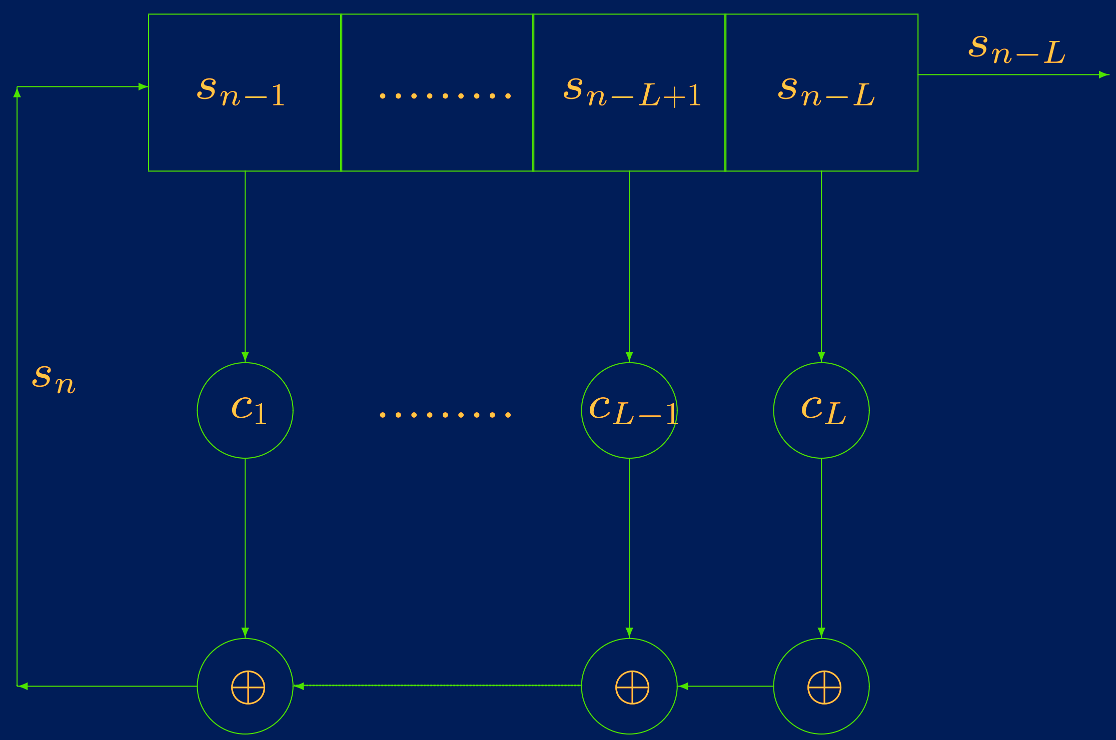
Jean-Marie Le Bars GREYC, U. de Caen, France

Alfredo Viola U. de la República, Uruguay

LIPN, U. Paris 13, France

STIC-AMSUD, November 22, 2007

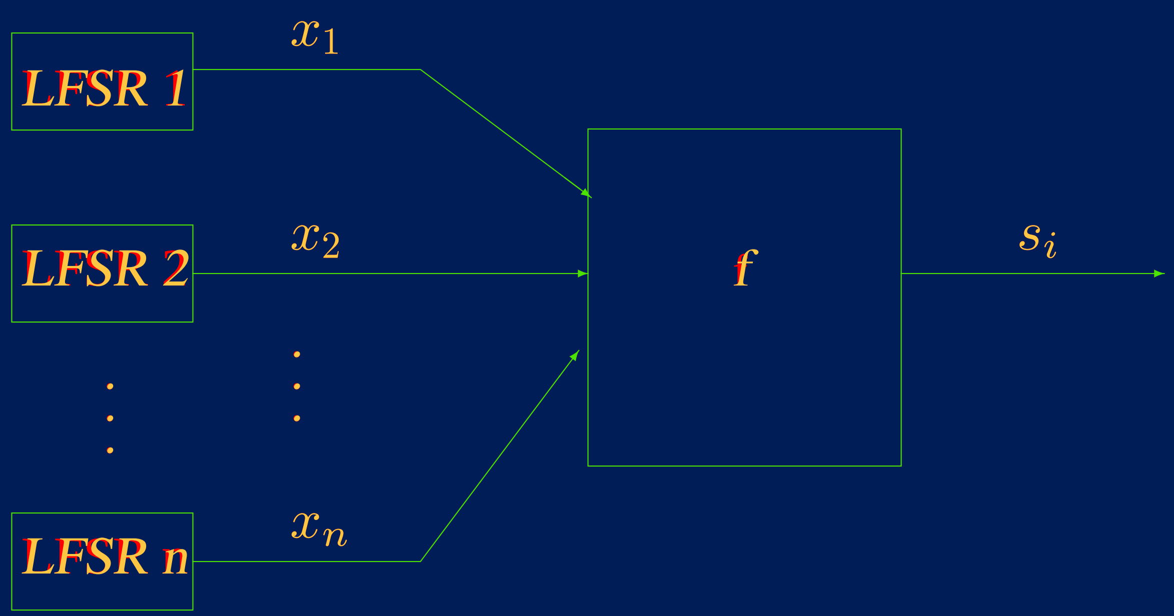
Linear Feedback Shift Registers (LFSR)



- At each clock-cycle computes $\bigoplus_{i=1}^{L} c_i s_{n-i}$ and outputs s_{n-L}.
- Generates an ultimately periodic sequence with period at most $2^L - 1$.
- The **linear complexity** of such a sequence is the length **L** of the minimum LFSR producing the same sequence.

Boolean functions and LFSR

- LFSR are cryptographically weak.
  - If L is the linear complexity of a sequence (unknown from the attacker), with **2L** consecutive bits known, the **Berlekamp-Massey** algorithm *recovers* **L, c_i's** and initial values (**key bits**).
  - In practice the attacker only needs to know around 20 consecutive bits.
- **Combining boolean functions** are used to avoid this attack.
- Period at most the LCM of the periods of the sequences generated by the LFSRs.
- Length of the key is $L_1 + L_2 + \ldots L_n$.



Cryptographic criteria for boolean functions

- High algebraic degree.
- Large Hamming distance to all affine functions.
- Balanced functions.
- Correlation-immune and resilient functions to prevent correlation attacks [Siegenthaler 1984; Meier & Staffelbach 1988; Johansson & Jönsson 1999, 2000; Canteaut & Trabbia 2000; and more ...].
- Strict avalanche criterion [Webster & Tavares 1985].
- Propagation criterion [Preneel, Van Leekwijck, Van Linden, Govaerts & Vandevalle 1991].
- [Carlet 2007] has a complete survey on boolean functions for cryptography and error correcting codes.
- Another useful source of information is [Gouguet 2004].

Decomposition of boolean functions

- A boolean function in n variables is a function $f : \{0,1\}^n \longrightarrow \{0,1\}$.
- We encode a boolean function f_n by a word in $\{0,1\}^{2^n}$ indexed by $x_n \ldots x_1$.
- We may see f_n as the concatenation of 2^{n-j} boolean functions in j variables each, by an operator we call $\star$.

Example:

f_1^1 f_1^2 f_1^3 f_1^4

f_3	*1*	*0*	*1*	*0*	*0*	*1*	*1*	*0*
x_3	0	0	0	0	1	1	1	1
x_2	0	0	1	1	0	0	1	1
x_1	0	1	0	1	0	1	0	1

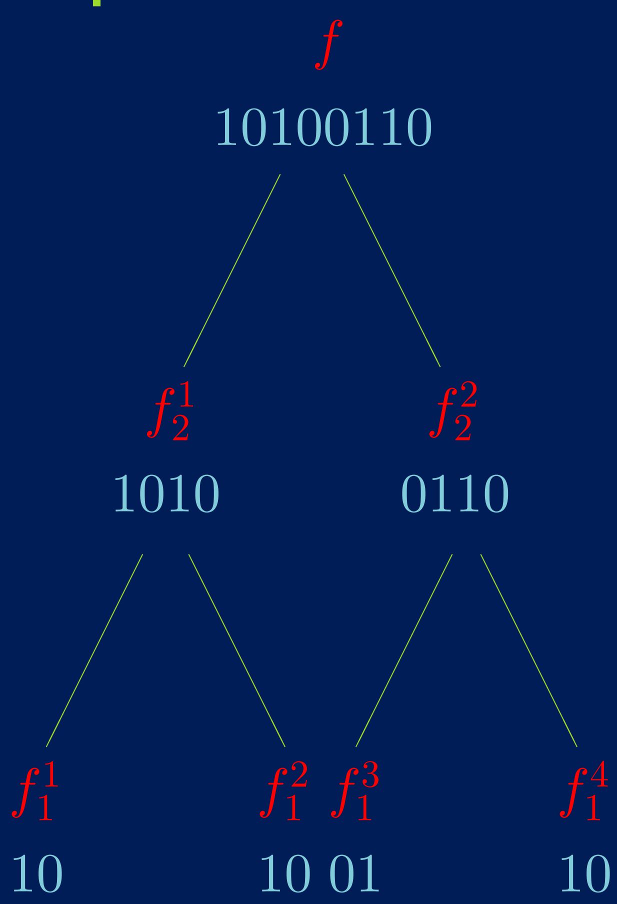
f_2^1 f_2^2

- We have $f_3 = f_2^1 \star f_2^2$ and $f_3 = f_1^1 \star f_1^2 \star f_1^3 \star f_1^4$.
- Functions f_j^k may be seen as f_n **conditioned** to some arguments x_i being some **fixed** a_i.
- Then, we have $f_1^3 = f_3 \mid_{x_2=0, x_3=1}$, $f_2^1 = f_3 \mid_{x_3=0}$ and so on.

Tree decomposition of a boolean function

- Decomposition can be described by a complete binary tree of depth $n-1$ being f_n is the root and the 2^{n-1} functions in 1 variable the leaves.

Example:



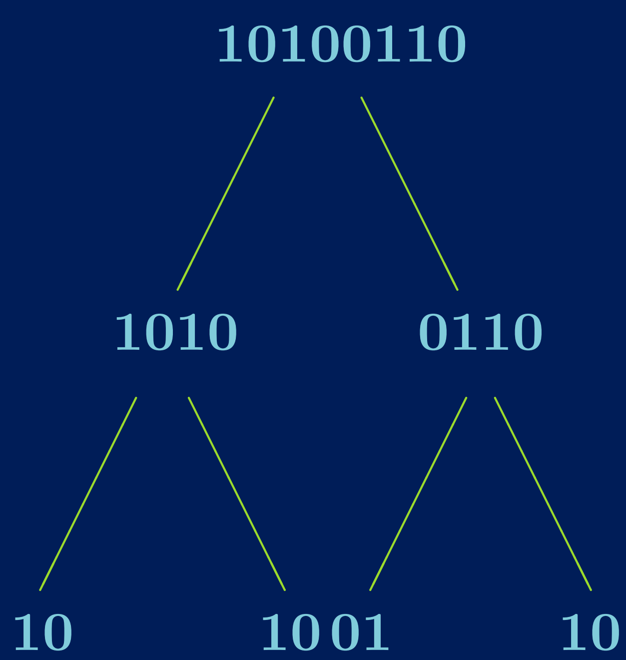
	f_1^1		f_1^2		f_1^3		f_1^4	
f_3	1	0	1	0	0	1	1	0
x_3	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**1**
x_2	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**1**
x_1	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**
	f_2^1				f_2^2			

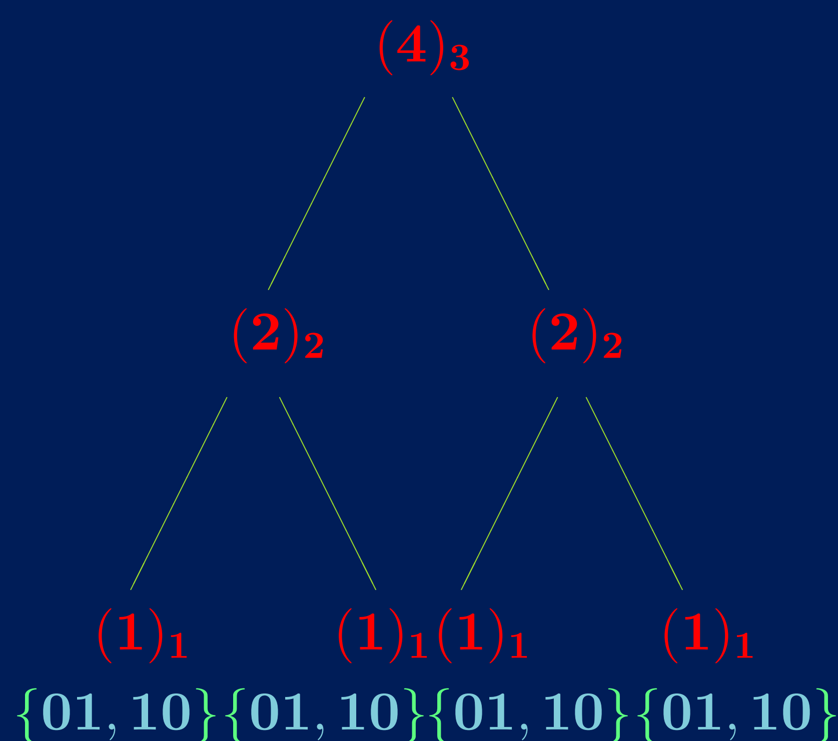
Hamming tree of boolean classes

- The sets $(i)_n = \{f_n \mid w_H(f_n) = i\}$ **partition** $\{0,1\}^{2^n}$ into **Hamming classes**.
- Similarly to boolean functions we have **Hamming trees** of boolean **classes**.
- Each tree may be **shared** by several functions.

Example:

- The Hamming tree of f_3 is shared by 16 functions.
- These functions are constructed by combining *01* and *10* in all possible ways at the leaves.



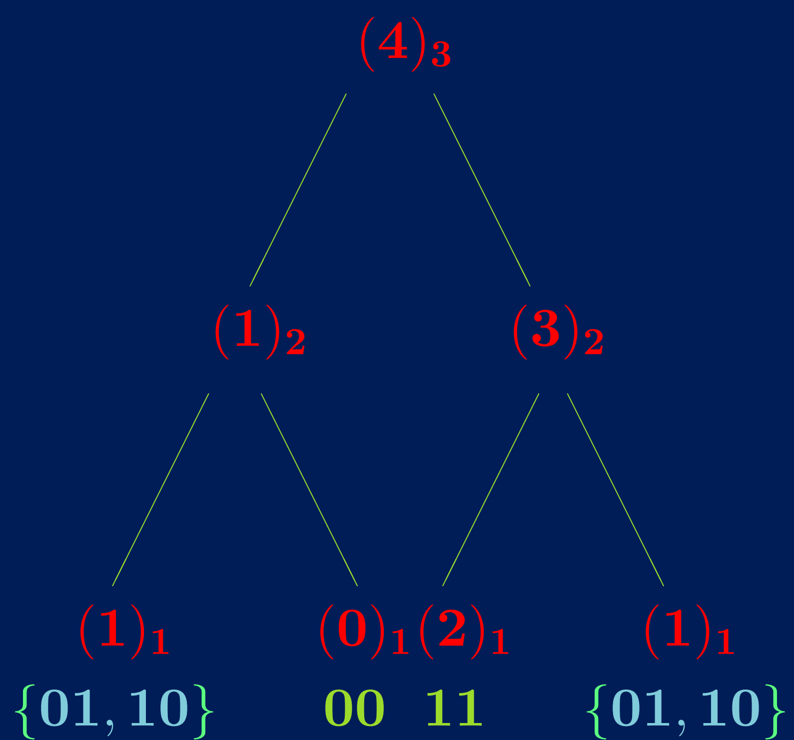


Hamming tree of boolean classes (cont.)

- Not every Hamming tree is shared by the same number of functions.

Example:

- This Hamming tree is shared by only 4 functions, since the class $(0)_1$ only contains the function 00 and $(2)_1$ only contains the function 11.
- These functions are 01001101, 01001110, 10001101 and 10001110.
- Nevertheless these functions are balanced like f_3 (belong to the same Hamming class), although their respective Hamming trees are different.
- **Hamming trees** (and not Hamming classes!) capture the **essential** features for our problem.



Equivalence relation for first-order correlation-immunity

- We define $\delta_i(f_n) = w_H(f_n \mid_{x_i=0}) - w_H(f_n \mid_{x_i=1}), 1 \leq i \leq n$.
- Then, f_n is **first-order correlation immune** $\iff \forall i, \delta_i(f_n) = 0$.
- Moreover, f_n is **1-resilient** $\iff \forall i, \delta_i(f_n) = 0, w_H(f_n) = 2^{n-1}$.
- A function f_n belongs to the **class** $\omega = \Omega(f_n) = \langle w_H(f_n), \delta_n(f_n) \ldots \delta_1(f_n) \rangle$.

How to find $\Omega(f_3)$

f_3	1	0	1	0	0	1	1	0
x_3	0	0	0	0	1	1	1	1
x_2	0	0	1	1	0	0	1	1
x_1	0	1	0	1	0	1	0	1

$\Omega(f_3) = \langle 4, \quad 0, \ , \ \rangle$.

How to find $\Omega(f_3)$

f_3	1	0	1	0	0	1	1	0
x_3	0	0	0	0	1	1	1	1
x_2	0	0	1	1	0	0	1	1
x_1	0	1	0	1	0	1	0	1

$$\Omega(f_3) = \langle 4, \quad 0, \ , \ \rangle.$$

f_3	1	0	1	0	0	1	1	0
x_3	0	0	0	0	1	1	1	1
x_2	0	0	1	1	0	0	1	1
x_1	0	1	0	1	0	1	0	1

$$\Omega(f_3) = \langle 4, \quad 0, 0, \ \rangle.$$

How to find $\Omega(f_3)$

f_3	1	0	1	0	0	1	1	0
x_3	0	0	0	0	1	1	1	1
x_2	0	0	1	1	0	0	1	1
x_1	0	1	0	1	0	1	0	1

$\Omega(f_3) = \langle 4, \quad 0, \ , \ \rangle$.

f_3	1	0	1	0	0	1	1	0
x_3	0	0	0	0	1	1	1	1
x_2	0	0	1	1	0	0	1	1
x_1	0	1	0	1	0	1	0	1

$\Omega(f_3) = \langle 4, \quad 0, 0, \ \rangle$.

f_3	1	0	1	0	0	1	1	0
x_3	0	0	0	0	1	1	1	1
x_2	0	0	1	1	0	0	1	1
x_1	0	1	0	1	0	1	0	1

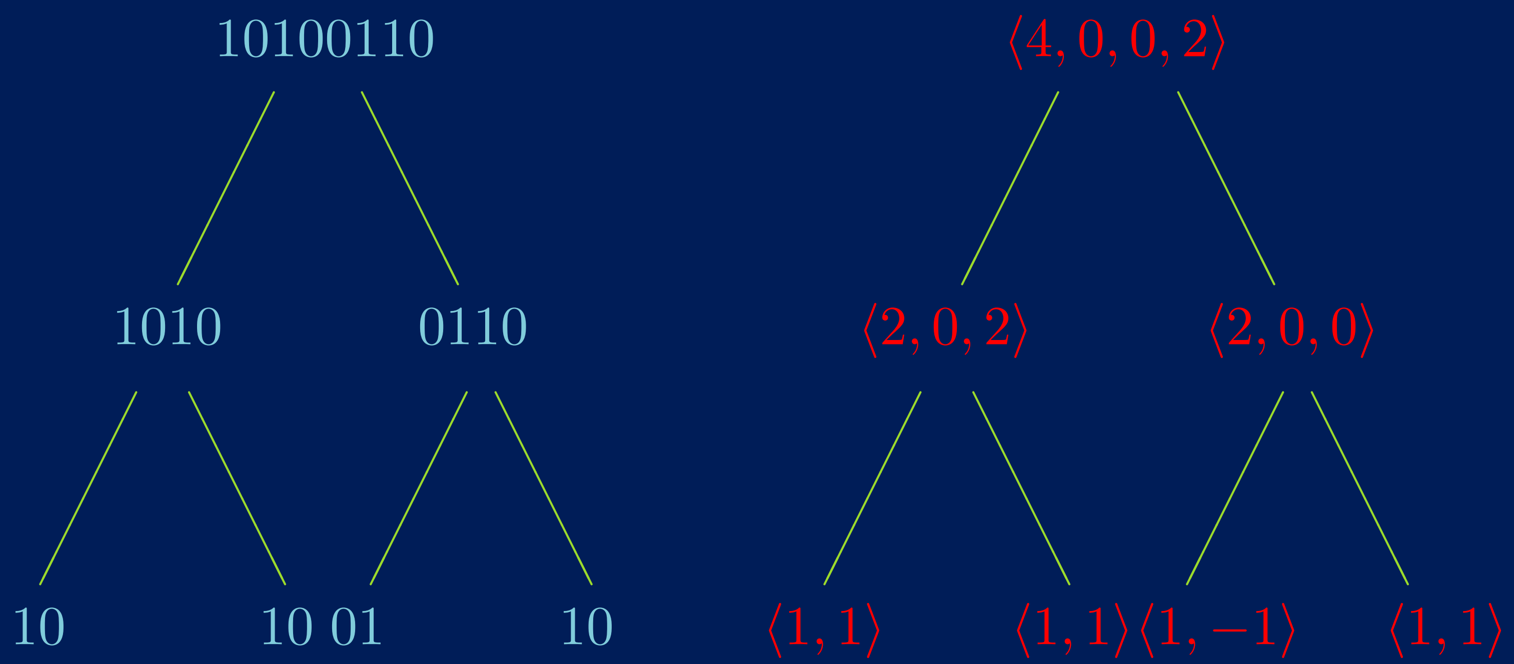
$\Omega(f_3) = \langle 4, \quad 0, 0, 2 \rangle$.

First order correlation-immune classes

- Similar to balanced functions we have **first order correlation trees**.

Example:

- This is the first order correlation tree for f_3.



Recursive construction of correlation classes

Proposition 1 (Recursive construction).

Let

$$\begin{cases} \omega_1 = \langle p_1, \delta^1_{n-1}, \ldots, \delta^1_1 \rangle \in \Omega^{p_1}_{n-1}, \\ \omega_2 = \langle p_2, \delta^2_{n-1}, \ldots, \delta^2_1 \rangle \in \Omega^{p_2}_{n-1}, \\ \omega = \langle m, \delta_n, \ldots, \delta_1 \rangle = \omega_1 \star \omega_2. \end{cases}$$

Then we have

$$\begin{cases} m = p_1 + p_2 \\ \delta_n = p_1 - p_2 \\ \delta_i = \delta^1_i + \delta^2_i, \quad i \in \{1, \ldots, n-1\}. \end{cases}$$

Decomposition of correlation classes.

Theorem 1 (Decomposition of correlation classes).

Let $\omega \in \Omega_n$. Then, with $\omega_1, \omega_2 \in \Omega_{n-1}$ we have

$$\omega^s = \bigcup_{\omega_1 \star \omega_2 = \omega} \omega_1^s \times \omega_2^s.$$

Theorem 2 (Counting correlation functions).

Let $\omega \in \Omega_n$. Then, with $\omega_1, \omega_2 \in \Omega_{n-1}$ we have

$$|\omega^s| = \sum_{\omega_1 \star \omega_2 = \omega} |\omega_1^s| \cdot |\omega_2^s|.$$

Decomposition of correlation-immune classes.

- Let $\omega_1 \in \Omega_{n-1}^{p_1}$, $\omega_2 \in \Omega_{n-1}^{p_2}$ and $m = p_1 + p_2$. From our recursive construction we have the equivalence $\omega_1 \star \omega_2 \in Cor_n^m \iff \omega_2 = \overline{\omega_1}$..

Theorem 3 (Decomposition of correlation-immune classes).

$$\mathcal{C}\mathbf{or}_{\mathbf{n}}^{\mathbf{m}} = \bigcup_{\omega_1 \in \mathbf{\Omega}_{\mathbf{n-1}}^{\mathbf{m}}} \omega_{\mathbf{1}}^{\mathbf{s}} \times \overline{\omega_{\mathbf{1}}}^{\mathbf{s}}, \text{ for } \mathbf{0} \leq \mathbf{m} \leq \mathbf{2}^{\mathbf{n-1}}.$$

Theorem 4 (Counting correlation-immune functions).

$$|\mathbf{Cor}_{\mathbf{n}}^{\mathbf{m}}| = \sum_{\omega_1 \in \mathbf{\Omega}_{\mathbf{n-1}}^{\mathbf{m}}} |\omega_{\mathbf{1}}^{\mathbf{s}}|^{\mathbf{2}},$$

$$|\mathbf{Cor}_{\mathbf{n}}| = \sum_{\omega_1 \in \mathbf{\Omega}_{\mathbf{n-1}}} |\omega_{\mathbf{1}}^{\mathbf{s}}|^{\mathbf{2}}.$$

Counting 1-resilient boolean functions.

- We denote by $\mathcal{B}_n$ the set of **balanced first-order correlation classes** with n variables.
- We have $|\mathcal{B}_n| = \binom{2^n}{2^{n-1}}$.

Corollary 4 (Counting 1-resilient boolean functions).

Since $Res1_n = Cor_n^{2^{n-2}}$, we have

$$|Res1_n| = \sum_{\omega_1 \in B_{n-1}} |\omega_1|^2.$$

- Then, to compute $Res1_n$ we only need to know the cardinality of all balanced first-order correlation classes with $n-1$ variables.
- We find an efficient algorithm by working with correlation classes and not with correlation functions.

n	*5*	*6*	*7*
$\mathrm{Res1}_n$	807980	95259103924394	23478015754788854439497622689296
Time	0.028 s	0.526 s	1 h 02 min 27.332 s

Normalized classes

- Let $m \leq 2^{n-1}$ and $\omega = \langle \mathbf{m}, \delta_{\mathbf{n}}, \ldots, \delta_{\mathbf{1}} \rangle \in \mathbf{\Omega}_{\mathbf{n}}^{\mathbf{m}}$. There exits a **permutation** $\sigma : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ which satisfies

$$\begin{cases} \alpha_i = |\delta_{\sigma(i)}| \\ \alpha_n \leq \alpha_{n-1} \leq \ldots \leq \alpha_1. \end{cases}$$

- The class $\mathbf{N}(\omega) = \theta = \langle \mathbf{m}, \alpha_{\mathbf{n}}, \ldots, \alpha_1 \rangle$ will be called the **normalised class** of ω.

- Every Boolean function in ω may be **transformed** in a **unique** Boolean function in θ.

Example:

- $N(\langle 7, 1, 5, -3, -3 \rangle) = \langle 7, 1, 3, 3, 5 \rangle$.

New characterization of 1-resilient functions.

- Each set ω^s corresponding to a class ω with $\mathbf{N}(\omega) = \theta$ has the **same cardinality as** θ^s. Then,

Theorem 5 (Number of 1-resilient functions).

$$|Res1_n| = \sum_{\theta \in \Theta_n^{2^{n-2}}} n(\theta)\ |\theta^s|^2.$$

- **Normalized classes** help to still **speed up** our counting algorithm and find the number of 1-resilient functions for $\mathbf{n = 7}$ in **50 seconds!**.

- By only computing a **fraction** of all normalized classes, we obtain the **lower bound** $4\ 10^{67}$ for the number of 1-resilient functions with $\mathbf{n = 8}$ variables.

Upper bounds on the number of first-order correlation classes

- Let $\omega = \langle m, \delta_n, \ldots, \delta_1 \rangle \in \Omega_n^m$. Then we define $\delta(\omega) = \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} |\delta_{\mathbf{i}}|$.
- We may see $\delta(\omega)$ as a measure of *how far from first-order correlation* is ω.
- Define $\delta(\mathbf{n}, \mathbf{m}) = \sup_{\omega \in \boldsymbol{\Omega}_{\mathbf{n}}^{\mathbf{m}}} \delta(\omega)$, $\mu(\mathbf{n}, \mathbf{m}) = \delta(\mathbf{n}, \mathbf{m})/\mathbf{2}$ if m is even and $\mu(\mathbf{n}, \mathbf{m}) = (\delta(\mathbf{n}, \mathbf{m}) - \mathbf{n})/\mathbf{2}$ otherwise.

Lemma 3 (Upper bound for number of classes)

Let $0 \leq m \leq 2^{n-1}$ and $\mathbf{U}_{\mathbf{n}}^{\mathbf{m}}$ defined by

$$U_n^m = \begin{cases} \sum_{j=0}^{n} \binom{\mu(n,m)}{j} \binom{\mu(n,m)+n-j}{\mu(n,m)} & m \text{ even} \\ \binom{\mu(n,m)+n}{n} 2^n & \text{otherwise.} \end{cases}$$

- U_n^m is an **upper bound** for the **number of classes in** Ω_n^m.

New lower bound on the number of 1-resilient functions

Theorem 7 $|Res1_n| \geq \frac{\binom{2^{n-1}}{2^{n-2}}^2}{U_{n-1}^{2^{n-2}}} \geq \frac{\binom{2^{n-1}}{m}^2}{\binom{\mu(n-1,2^{n-2})+n-1}{n-1}2^{n-1}}.$

Theorem 8 $|Res1_n| \geq \frac{2^{2^n}(n\pi)^{n/2}}{2^{n^2-\frac{3}{2}n-1}e^{n-1/2}}.$

- [Maitra & Sarkar 1999]:

$$|Res1_n| \geq 2^{2^{n-2}} + \binom{2^{n-1}}{2^{n-2}} + \binom{2^{n-2}}{2^{n-3}} * \left(\binom{2^{n-2}}{2^{n-3}} - 2\right) + \binom{2^{n-3}}{2^{n-4}} - 2^{2^{n-3}}.$$

n	*[Maitra & Sarkar 1999]*	*Our lower bound*
5	*17876*	**503430**
6	$7.667\ 10^{8}$	$\mathbf{7.523\ 10^{12}}$
7	$2.193\ 10^{18}$	$\mathbf{1.312\ 10^{29}}$
8	$2.730\ 10^{37}$	$\mathbf{1.134\ 10^{64}}$
9	$6.342\ 10^{75}$	$\mathbf{8.884\ 10^{136}}$
10	$5.058\ 10^{152}$	$\mathbf{2.128\ 10^{286}}$

- Dramatic improvements are due to our **general construction**. The previous bounds have been found by building and counting **restricted classes**.

New lower bound on the number of k-resilient functions

- Given a 1-resilient function in $n-k+1$ variables, we have a construction that leads to a k-resilient function in n variables.
- As a consequence we have the following new lower bound for the number of k-resilient functions:

Theorem 9 Let $k \geq 2$, and $n > k$. The set of k-resilient functions with n variables satisfies $|\mathbf{Res^1_{n-k+1}}| \leq |\mathbf{Res^k_n}|$.

n	*Maiorana-McFarland* [Camion,Carlet,Charpin & Sendrier 1991]	*Our lower bound*
$\mathcal{R}es^1_{10}$	$3.0\ 10^{79}$	$\mathbf{5.1\ 10^{285}}$
$\mathcal{R}es^2_{10}$	$4.3\ 10^{40}$	$\mathbf{3.4\ 10^{136}}$
$\mathcal{R}es^3_{10}$	$1.2\ 10^{21}$	$\mathbf{2.6\ 10^{63}}$
$\mathcal{R}es^4_{10}$	$1.4\ 10^{11}$	$\mathbf{2.3\ 10^{31}}$
$\mathcal{R}es^5_{10}$	$1.1\ 10^{6}$	$\mathbf{9.5\ 10^{13}}$

Summary of results

- We present a complete characterization of 1-correlation immune functions and give efficient algorithms to generate and count them.
- The number of 1-resilient functions in $\mathbf{7}$ variables is $\mathbf{23478015754788854439497622689296}$.
- We drastically improve knwon bounds specially lower bonds.

n	*8*	*9*	*10*	*11*	*12*	*13*
Maitra	10^{37}	10^{75}	10^{152}	10^{306}	10^{614}	10^{1231}
New Lower Bound	$\mathbf{10^{64}}$	$\mathbf{10^{136}}$	$\mathbf{10^{286}}$	$\mathbf{10^{589}}$	$\mathbf{10^{1199}}$	$\mathbf{10^{2426}}$
New Upper Bound	$\mathbf{10^{68}}$	$\mathbf{10^{144}}$	$\mathbf{10^{297}}$	$\mathbf{10^{603}}$	$\mathbf{10^{1218}}$	$\mathbf{10^{2449}}$
Schneider	10^{71}	10^{147}	10^{299}	10^{606}	10^{1221}	10^{2452}

- We conjecture that the *probability* of a boolean function being 1-*resilient* is

$$\sim \frac{\mathbf{2}^{-\frac{\mathbf{n}^2}{\mathbf{2}}}}{\sqrt{\mathbf{2}\pi}^{\frac{\mathbf{n}+\mathbf{1}}{\mathbf{2}}}}.$$

- Use of the generating function derived from our constructions. Work in progress with P. Flajolet, S. Mesnager.