

Correlation of Boolean Functions

Guang Gong
Department of Electrical & Computer
Engineering
University of Waterloo, CANADA

Presentation Outline

- Sequences, Correlation, Cryptographic Properties, Cryptanalysis, and Their Relation to Transforms for Signals
- Indication Functions: A Bridge to Connect Resiliency (Cross Correlation) and Propagation (Additive Autocorrelation)
- Constructions of Boolean Functions with 2-Level (Multiplicative) AC and Three-valued Additive AC, and more
- Discussions

Applications of Pseudo-random Sequences

In communications:

- Orthogonal codes, cyclic codes
- CDMA (code division multiple access) applications
- Synchronization codes
- Radar, and deep water distance range
- Testing vectors of hardware design
-

In cryptography:

- Key Stream Generators in Stream Cipher Models
- Functions in Block Ciphers
- Session Key Generators
- Pseudo-random Number Generators in Digital Signature Standard (DSS), etc.
- Digital Water-mark
-



Design of Pseudo-random Sequence Generators

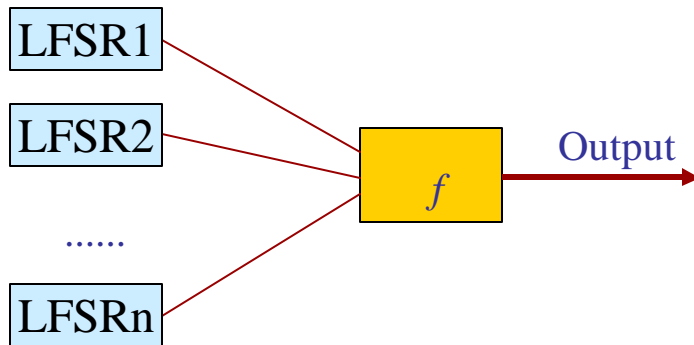
(a) Towards 2-Level
Auto-Correlation
and Low Correlation

(b) Towards Large
Linear Span

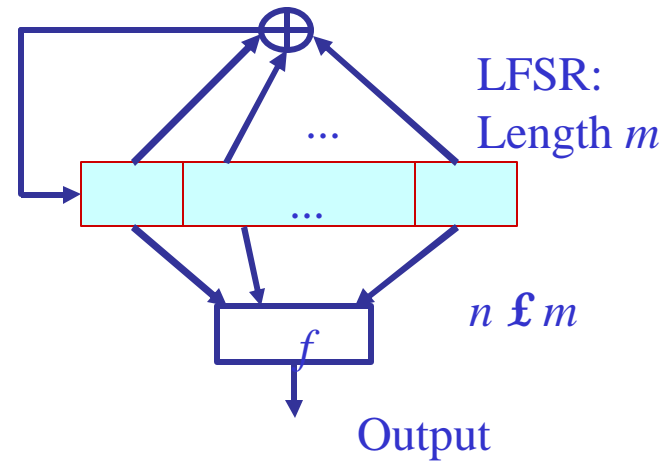


LFSR as Basic Blocks

Stream Cipher Applications



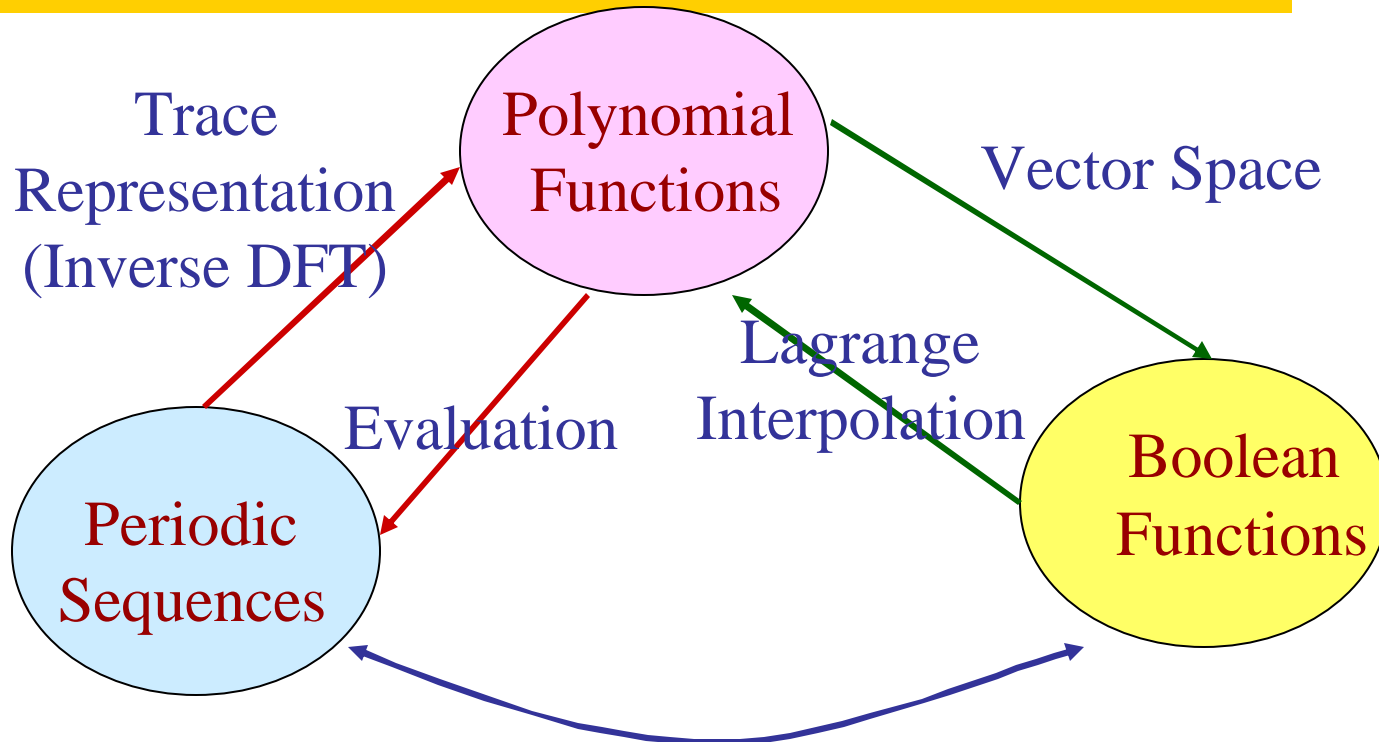
A Combinatorial Function Generator



A Filtering Generator

f is a boolean function in n variables .

1-1 Correspondences Between Sequences, Polynomial Functions and Boolean Functions



Notation



➤ $F = \text{GF}(2^n)$, a finite field, \mathbf{a} is a primitive element of F
 $F_2 = \text{GF}(2)$, binary field.

➤ $\mathbf{a} = \{a_i\}$, a binary sequence with period $N / 2^n - 1$; $f(x)$,
the trace representation of \mathbf{a} , i.e.,

$$a_i = f(\mathbf{a}^i), \quad i = 0, 1, \dots$$

Note. $f(x)$ is a polynomial function from $\text{GF}(2^n)$ to $\text{GF}(2)$
which can be represented by

$$f(x) = \sum_k \text{Tr}_1^{n_k}(A_k x^k), \quad A_k \in \text{GF}(2^{n_k})$$

where the k 's are different coset leaders modulo $2^n - 1$, and n_k is
the size of the coset containing k .

➤ $x = x_0 + x_1 \mathbf{a} + \dots + x_{n-1} \mathbf{a}^{n-1} = (x_1, \dots, x_n)$, an element in finite field
 $\text{GF}(2^n)$ or an element in the vector space F_2^n .

(Multiplicative) Autocorrelation

The (multiplicative) autocorrelation of function $f(x)$ is defined as the autocorrelation of the sequence \mathbf{a} , which is given by

$$C_f(\mathbf{t}) = 1 + C(\mathbf{t}) = 1 + \sum_{i=0}^{N-1} (-1)^{a_{i+t} + a_i}, \quad \mathbf{t} = 0, 1, \dots$$

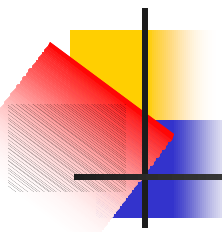
The sequence \mathbf{a} has an (ideal) 2-level autocorrelation if

$$C(\mathbf{t}) = \begin{cases} N & \text{if } \mathbf{t} \equiv 0 \pmod{N} \\ -1 & \text{otherwise} \end{cases}$$

Convolution or Additive Autocorrelation

The additive autocorrelation of f (or the additive autocorrelation of the sequence is defined through its trace representation f) is defined as the convolution of $f(x)$:

$$A_f(w) = \sum_{x \in F} (-1)^{f(x)+f(x+w)}$$



Known Constructions of 2-Level Autocorrelation Sequences (or Orthogonal Codes, or Hadamard Difference Sets)

- Number theory approach (N is a prime): quadratic residue sequences (with $N \equiv 3 \pmod{4}$), Hall sextic residue sequences, and the twin prime sequences.

$$N = 2^n - 1:$$

- PN-sequences = m -sequences (1931, Singer, 1958, Golomb)
- GMW sequences (1961, Golden-Miller-Welch, 1984, Welch-Scholtz)
- Conjectured Sequences (Gong-Gaal-Golomb, 1997, No-Golomb-Gong-Lee-Gaal, 1998)
- Hyper-oval Construction: (Maschietti, 1998)
- Kasami Power Function Construction (Dobbertin, Dillon, 1998)

2-level Additive Autocorrelation

$f(x)$ is a bent function if and only if

$$\hat{f}(\mathbf{1}) = \pm\sqrt{2^n}, \quad \forall \mathbf{1} \in F$$

Note. Bent functions only exists for n even.

$f(x)$ has 2-level additive autocorrelation if and only if $f(x)$ is bent. There are two general constructions for bent functions (compared with the constructions of the binary sequences with 2-level (multi.) autocorrelation, this is relatively easy).

Question: What is the best additive autocorrelation for n odd?

Transforms for Signal (Sequence) Design (Engineering Perspective)

Hadamard (Walsh)
Transform of f :

$$\hat{f}(I) = \sum_{x \in F} (-1)^{\text{Tr}(Ix) + f(x)}$$

Time
domain

$f(x)$

Frequency
domain

$\hat{f}(I)$

Convolution or
Additive
autocorrelation of f :

$$A_f(w) = \sum_{x \in F} (-1)^{f(x) + f(x+w)}$$

They are related by the Convolution Law.

In other words, the square of the Hadamard transform of f is equal to the Hadamard transform of the convolution of f with itself or additive autocorrelation of f . Conversely,

$$A_f(w) = \frac{1}{2^n} \sum_{I \in F} (-1)^{\text{Tr}(wI)} \hat{f}^2(I)$$

which is a fundamental relation through this representation.

Desired Cryptographic Properties of Boolean Functions



Definition 1 (Siegenthaler, 1984)

A Boolean function $f(x)$ in n variables is **k th-order correlation immune** if for each k -subset K of $\{0, \dots, n-1\}$, $Z = f(x)$, considered as a random variable over F_2 , is independent of all x_i for $i \in K$. Furthermore, if $f(x)$ is balanced and k th-order correlation immune, then $f(x)$ is said to be **k -order resilient**.

Nonlinearity of f is defined as the minimum distance of $f(x)$ with all affine functions, or equivalently,

$$N_f = 2^{n-1} - \frac{1}{2} \max_I |\hat{f}(I)|$$

Property (Xiao and Massey, 1988). $f(x)$ is **k th-order correlation immune** if and only if

$$\hat{f}(I) = 0, 1 \leq H(I) \leq k$$

where $H(x)$ is the Hamming weight of x .

A historical remark.

Golomb studied these concepts under the terminology of **invariants** of boolean functions in 1959, and he is the first to compute them using Hadamard transform.

Definition 2.

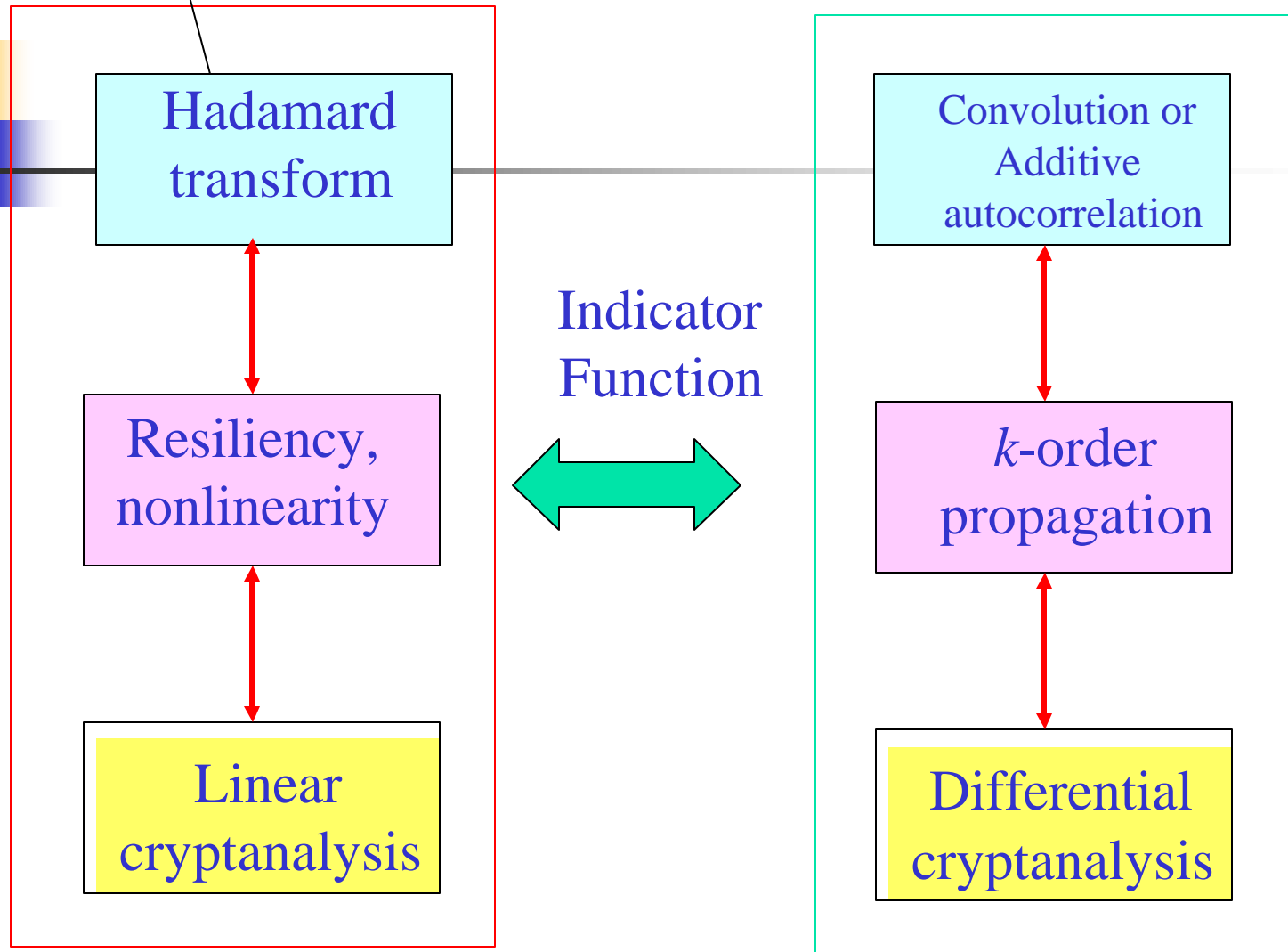
A Boolean function $f(x)$ in n variables is said to satisfy the avalanche criterion (SAC) if

$$A_f(w) = 0 \quad \text{for all } w \text{ with } H(w) = 1$$

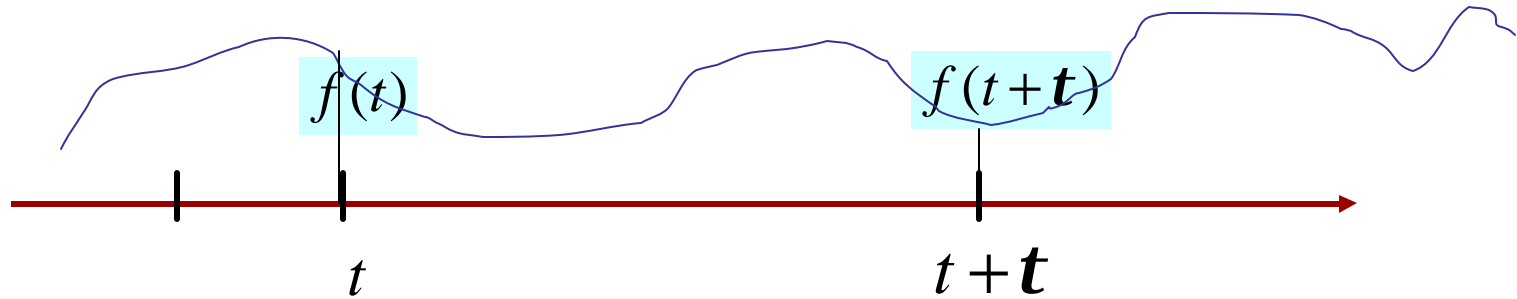
to have the k -order propagation if

$$A_f(w) = 0 \quad \text{for all } w \text{ with } 1 \leq H(w) \leq k$$

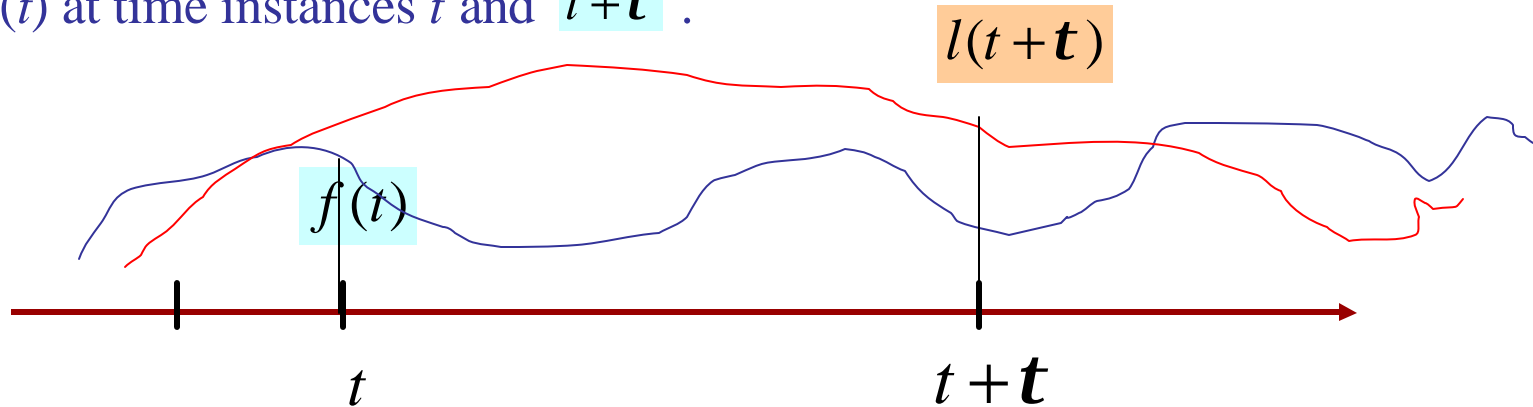
(cross correlation with m-sequences)



Engineering Perspective of Differential Cryptanalysis and Linear Cryptanalysis Associated to Transforms for Signal (Sequence) Design



Differential cryptanalysis (or propagation) is to exploit the correlation of the signal $f(t)$ at time instances t and $t+t$.



Correlation immunity (or resiliency, nonlinearity, linear cryptanalysis) is to exploit the correlation between the signal $f(t)$ and the reference linear signal $l(t)$ at time instances t and $t+t$.

Indicator Function: A Bridge for Connecting Resiliency and Additive Autocorrelation

Definition. An indicator function of f , denoted by $\mathbf{s}_f(x)$, is defined as

$$\mathbf{s}_f(\mathbf{1}) = \begin{cases} 0 & \text{if } \hat{f}(\mathbf{1}) = 0 \\ 1 & \text{if } \hat{f}(\mathbf{1}) \neq 0 \end{cases}$$

Example. For $n = 5$, $\text{GF}(2^5)$ defined by $\mathbf{a}^5 + \mathbf{a}^3 + 1 = 0$, and

$$f(x) = \text{Tr}(x^3) .$$

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
$\hat{f}(\mathbf{a}^i)$	-8	0	0	0	0	-8	0	8	0	-8	-8	8	0	8	8	0	0	0	-8	8	-8	8	8	0	0	8	8	0	8	0	0
$\mathbf{s}_f(\mathbf{a}^i)$	1	0	0	0	0	1	0	1	0	1	1	1	0	1	1	0	0	0	1	1	1	1	1	0	0	1	1	0	1	0	0

Preferred set: For $n = 2m + 1$, f is said to be preferred if the Hadamard transform of f has the following three values:

$$P = \{0, \pm 2^{m+1}\}$$

Optimal Additive autocorrelation (AC): For $n = 2m + 1$, let f be balanced, the additive AC of f is said to be optimal if the maximal magnitude of the additive AC at nonzero, denoted as

Δ_f , is 2^{m+1} and A_f has 2^{n-1} zeros in $\text{GF}(2^n)$.

Note.

1. According to the Parseval energy formula, 2^{m+1} is minimum among magnitudes of all 3-valued Hadamard spectra.
2. Zhang and Zheng (1995) conjectured that*

$$\Delta_f \geq 2^{m+1}$$

Observation 1: Indicator Function and Resiliency

Let f be preferred. Then f is 1-order resilient if and only if the dual of f is nonlinear.

Zhang and Zheng, 1999 under boolean forms, Gong and Youssef, 1999 under polynomial forms, Canteaut-Carlet-Charpin-Fontaine, 2000, for any three-valued Hadamard transform.



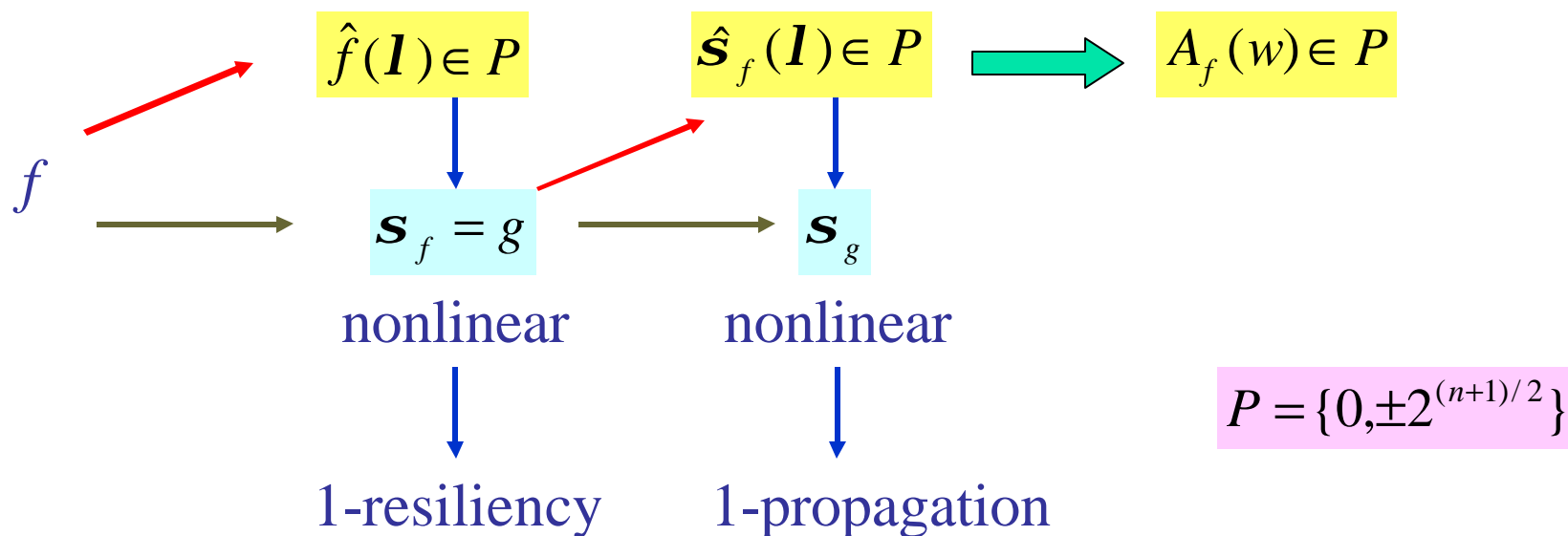
Observation 2: Indicator Function and Additive Autocorrelation

Let f be preferred. Then the additive autocorrelation functions of f at nonzero is equal to opposite of the Hadamard transform of f . In other words,

$$A_f(w) = -\hat{S}_f(w), \quad \forall 0 \neq w \in \text{GF}(2^n)$$

Theorem. A Sufficient Condition for Preferred Additive Autocorrelation

If the Hadamard transforms of both f and its indicator functions are preferred, then the additive autocorrelation is preferred.



Constructions

All functions, listed in Tables 1-3, are from binary sequences with 2-level (multiplicative) autocorrelation.

Cryptographic Properties:

- a) 2-level AC
- b) Nonlinearity: $2^{n-1} - 2^{(n-1)/2}$
- c) Preferred f
- d) 1-order resiliency
- e) Preferred additive AC, so optimal additive AC
- f) 1-order propagation.

Table 1. Properties (a)-(d)

Functions from the sequence sets	Indicator Functions	Comments
Kasami decimation: $Tr(x^d), d = 2^{2k} - 2^k + 1$	$Tr(x^{2^k+1}), \text{ for } 3k \equiv 1 \pmod n$	Kasami 1971, Dillon 1999
The other Kasami, Welch, Niho	nonlinear	
Subset of GMW sequences	Nonlinear	2-level AC (Goldon, Miller, Welch 1961) HT (Games (85), Klapper(96))
Welch-Gong sequences $WG(x)$	$Tr(x^{d^{-1}})$	2-level AC (No <i>et. al</i> 1998, Dillon <i>et. al.</i> 1999)
Glynn Type 1 hyperoval sequences	$Tr(x^{(k-1)/k})$	2-level AC (Matchietti 1998), Hadamard transform (Xiang 1998, Dillon 1999)
Kasami power function sequences: $C_k(x)$	$Tr(x^{(2^k+1)/3})$	2-level AC (No <i>et. al</i> 1998, Dillon <i>et. al.</i> 1999)

Table 2. Properties (a)-(e)

$$3k \equiv 1 \pmod{n}$$

(Boolean) Functions	Indicator Functions
Kasami sequences: $Tr(x^d), d = 2^{2k} - 2^k + 1$	$Tr(x^{2^k+1})$
Welch-Gong sequences $WG(x) = Tr(t(x+1)+1)$	$Tr(x^{d^{-1}})$
Kasami power function sequences: $C_3(x), k = 3$ $C_k(x) = Tr(t(x^{2^k+1}))$	$Tr(x^3)$ $Tr(x^{d^{-1}})$

where

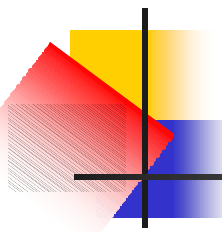
$$t(x) = x + x^{2^k+1} + x^{2^{2k}+2^k+1} + x^{2^{2k}-2^k+1} + x^{2^{2k}+2^k-1}$$

Table 3. Properties (a)-(f)

(Boolean) Functions	Indicator Functions
Welch-Gong sequences $WG(x) = Tr(t(x+1) + 1)$	$Tr(x^{d^{-1}})$
Kasami power function sequences (5-term sequences): $C_k(x) = Tr(t(x^{2^k+1}))$	$Tr(x^{d^{-1}})$

where

$$t(x) = x + x^{2^k+1} + x^{2^{2k}+2^k+1} + x^{2^{2k}-2^k+1} + x^{2^{2k}+2^k-1}, 3k \equiv 1 \pmod{n}$$



Example 11.7 Let $n = 7$. Then $k = 5 \implies n - k = 2 \implies 2^{n-k} + 1 = 5$, and $t(x) = x + x^5 + x^{21} + x^{13} + x^{29}$. Thus

$$\begin{aligned}C_5(x) &= \text{Tr}(t(x^{2^2+1})) = \text{Tr}(x^5 + x^{19} + x^{29} + x^3 + x^9) \\WG(x) &= \text{Tr}(t(x+1) + 1) = \text{Tr}(x + x^3 + x^7 + x^{19} + x^{29}).\end{aligned}$$

Both $C_5(x)$ and $WG(x)$ have the following properties:

- (a) Orthogonal or 2-level autocorrelation.
- (b) Nonlinearity $N_f = 56$.
- (c) Hadamard transform is preferred, i.e., belongs to $\{0, \pm 16\}$.
- (d) 1-resiliency under some basis.
- (e) The additive autocorrelation function is preferred, i.e., belongs to $\{0, \pm 16\}$.
- (f) 1-order propagation under some basis.

Discussions

- What are the additive autocorrelations of the rest functions with 2-level autocorrelation?
- The functions constructed from sequence design do not have linear structure for any fixed set of input variables (possible week leakage of Maiorana-McFarland like resilient functions).
- Experimental results show that there are many functions having preferred Hadamard transform, and preferred additive AC, so optimal additive AC, but not 2-level AC.



References

G. Gong and K.M. Khoo,

Additive autocorrelation of resilient boolean functions,

Proceedings of Tenth Annual Workshop on Selected Areas in
Cryptography (SAC), August 11-12, 2003, Ottawa, Canada, *Lecture
Notes in Computer Science*, Springer-Verlag, 2003.

Welcome to Waterloo for SAC 2004!