On Ideal t-Tuple Distribution of Orthogonal Functions in Filtering de Bruijn Generators

Guang Gong

Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, ON, N2L 3G1, CANADA

MMC Workshop and Tor Helleseth Symposium, Lofoten, Norway, Sep. 4-9 , 2017

This is a joint work with Kalikinkar Mandal.

UNIVERSITY OF WATERLOO

Tor!

Outline

- Introduction to randomness generation
- Basic concepts and properties
- Invariants under the WG transform
- Ideal tuple distribution of WG transformations in filtering de Bruijn sequences
- Concluding remarks and some open problems

How to generate ideal distributed random bits?

- For both linear feedback shift register (LFSR) sequences of with primitive polynomials of period $2^n - 1$ and de Bruijn sequences of period 2^n, each n-tuple occurs (excluded zero n-tuple for the LFSRs' case) exactly once in one period!

Given a bit stream

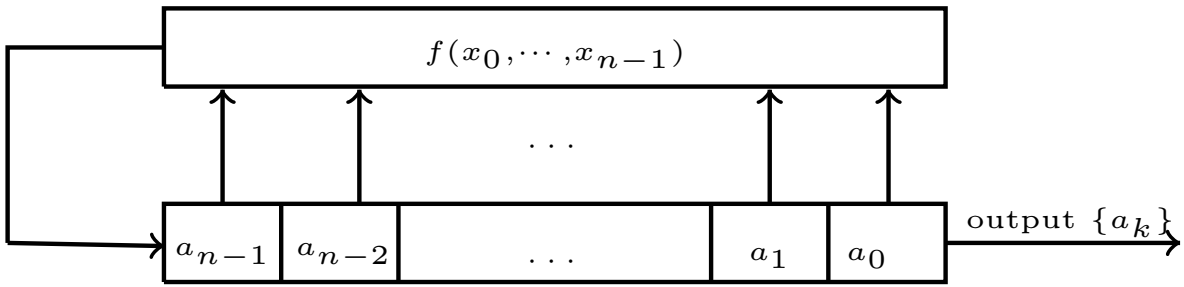
0000100110101111

with period 16, what is the distribution of a 4-tuple?

0000	1010
0001	0101
0010	1011
0100	0111
1001	1111
0011	1110
0110	1100
1101	1000

Each 4-bit pattern (e.g., 4-tuple) occurs exactly once in one period of the given sequence!

Nonlinear Feedback shift register (NLFSR) sequences

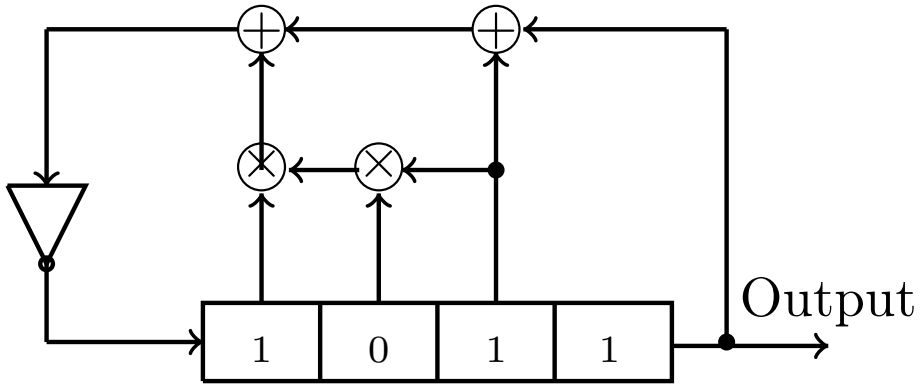


The feedback function is a boolean function in n variables:

$$f(x_0, x_1, \cdots, x_{n-1}) = \sum c_{i_1 i_2 \cdots i_t} x_{i_1} x_{i_2} \cdots x_{i_t}, c_{i_1 i_2 \cdots i_t} \in \mathbb{F}_2$$

where the sum runs through all subsets $\{i_1, \cdots, i_t\}$ of $\{0, 1, \cdots, n-1\}$.

Example of 4-stage NLFSR



Nonlinear feedback function	$f(x_0, x_1, x_2, x_3) = x_0 + x_1 + x_1x_2x_3 + 1$
The initial state	$(a_3, a_2, a_1, a_0) = 1011$
The output sequence	$a_0, a_1, \cdots =$ $1101100101000011\cdots$
Period	16

This generates a de Bruijn sequence of period 16!

De Bruijn Sequences (de Bruijn, 1946)

- A binary de Bruijn sequence of period 2^n is an output sequence of an nonlinear feedback shift register (NLFSR) of order n, i.e., each n-tuple occurs exactly once in one period of the sequence.
- **Open problem (Golomb 1967) on NLFSR analysis**: how to select the boolean function f such that the output has the full period 2^n?

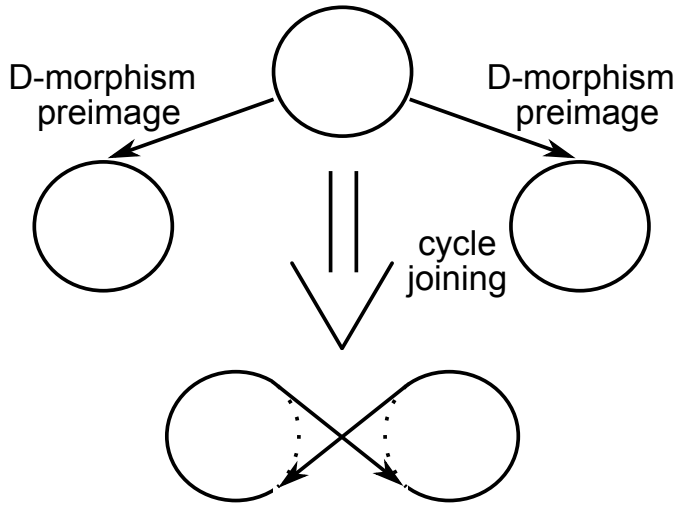
Known Constructions

- Composited construction using D-homomorphism (Lempel (70), Mykkletveit (79))
- Joining cycles (Fredricksen (82), Dai *et al.* (88), $\cdots$, Li *et al.* (2015), $\cdots$)

Composited Construction

To construct a de Bruijn sequence of period 2^{n+1} from a de Bruijn sequence of period 2^n,

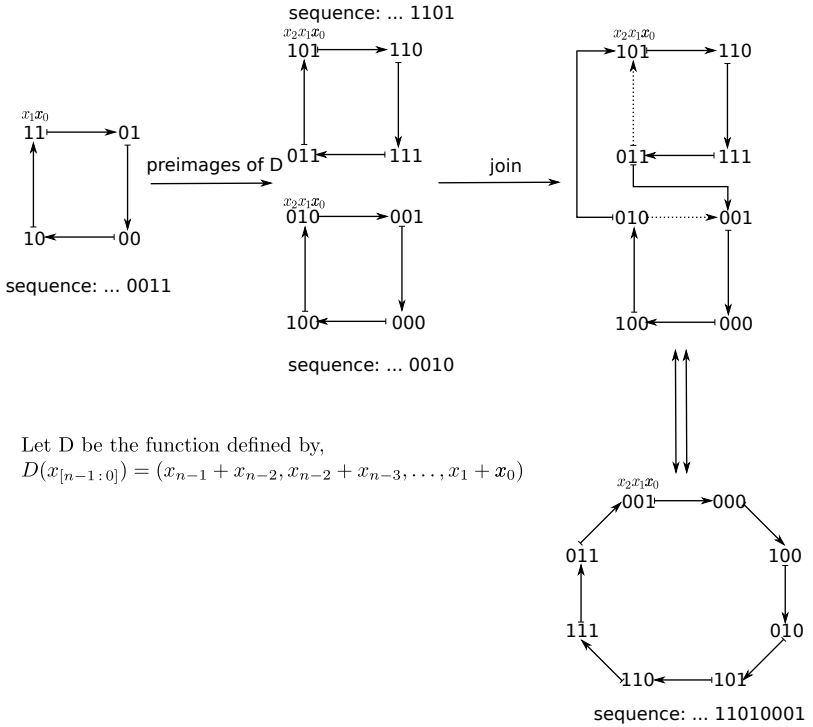
- compute two D-morphic preimages of the de Bruijn sequence of period 2^n,
- join(concatenate) these two preimages at a conjugate pair.



Example of Composited Construction

- For de Bruijn sequence of period 4. The two D-homomorphic preimage is given by the second row of each array below.
- The process for getting a de Brujn sequence of period 8 is given on the right figure.

	0	0	1	1
1	1	1	0	1
	0	0	1	1
0	0	0	1	0



Randomness Measurements

- ***Run:*** For a binary sequence **a** with period N, k consecutive zeroes (or ones) preceded by one (or zero) and followed by one (or zero) is called a *run of zeroes (or ones) of length k*. E.g.,

  $$\underline{0000}1001\underline{11}0101111$$

- ***Autocorrelation and crosscorrelation:*** Let **a** and **b** be two binary sequences with period N,

$$C_{\mathbf{a},\mathbf{b}}(\tau) = \begin{cases} \sum_{i=0}^{N-1}(-1)^{a_i+a_{i+\tau}} & \textbf{autocorrelation function of a if a = b} \\ \sum_{i=0}^{N-1}(-1)^{a_i+b_{i+\tau}} & \textbf{crosscorrelation function of a if b is not a cyclic shift of b} \end{cases}$$

Example

Let $\mathbf{a} = 1001011$ and $\mathbf{b} = 1110100$, both are m-sequences of period 7.

$$C_{\mathbf{x}}(\tau) = \begin{cases} 7 & \text{for } \tau \equiv 0 \pmod 7 \\ -1 & \text{for } \tau \not\equiv 0 \pmod 7 \end{cases}$$

for both $\mathbf{x} \in \{\mathbf{a}, \mathbf{b}\}$.

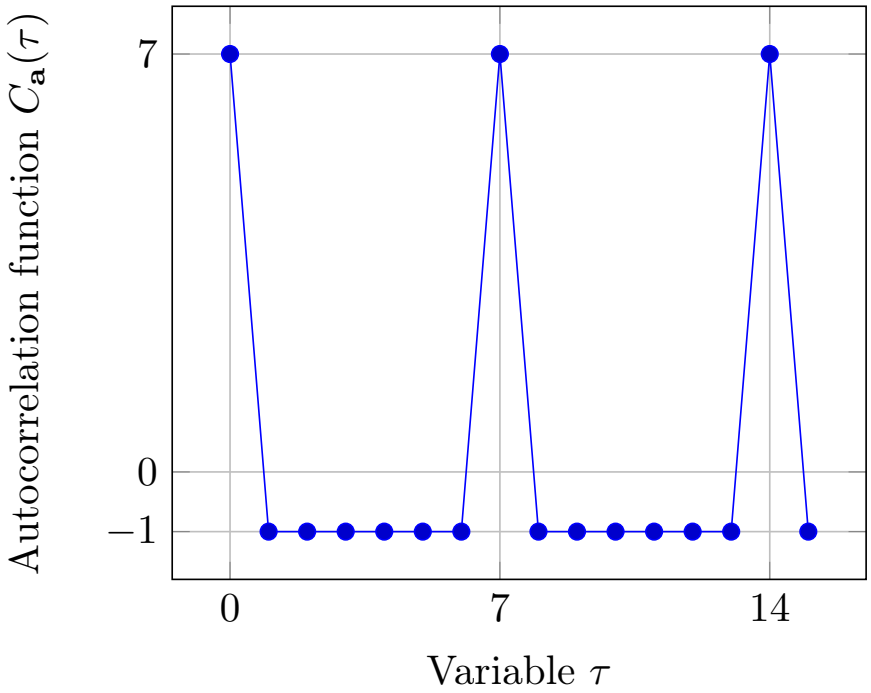


Figure: Autocorrelation of $\mathbf{a}$ or $\mathbf{b}$ which are 2-leved

Example (cont.)

τ	0	1	2	3	4	5	6
$C_{\mathbf{a},\mathbf{b}}(\tau)$	-5	3	3	-1	3	-1	-1

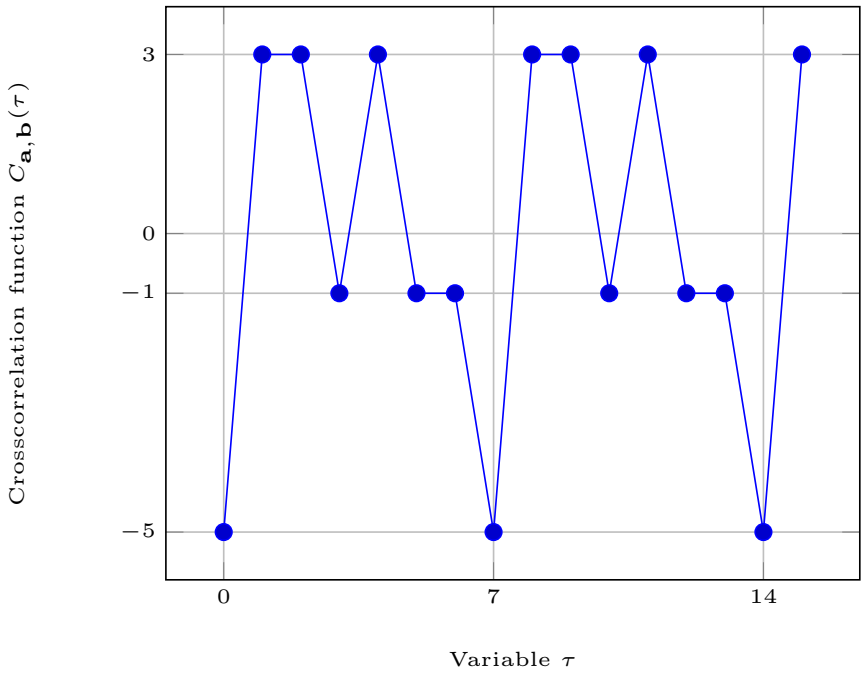


Figure: Crosscorrelation of 1001011 and 1110100

Golomb's Three Randomness Postulates for Binary Sequences

- **R-1**. In every period, the number of zeroes is nearly equal to the number of ones, i.e., the disparity is not to exceed 1.
- **R-2**. In every period, **half the runs** have length one, one-fourth have length two, one-eighth have length three, etc., as long as the number of runs so indicted exceeds 1. Moreover, for each of these lengths, there are equally many runs of 0's and of 1's.
- **R-3**. The autocorrelation function $C(\tau)$ is two-valued, given by

$$C(\tau) = \begin{cases} N & \textbf{if } \tau \equiv 0 \pmod{N} \\ K & \textbf{if } \tau \not\equiv 0 \pmod{N} \end{cases} \tag{1}$$

where K is a constant. If $K = -1$ for N odd and $K = 0$ for N even, then we say that the sequence has the *(ideal) 2-level autocorrelation function*.

Ideal tuple distribution of binary sequences

Ideal t-tuple distribution or uniformity of t-tuple distribution

For a binary sequence $\mathbf{a} = \{a_i\}$ of period N. We say the sequence $\mathbf{a}$ has *t-tuple distribution* if every t-tuple $(a_i, a_{i+1}, \cdots, a_{i+t-1})$ occurs equally likely in one period of $\mathbf{a}$. In particular, if $n = 2^n$, each t-tuple occurs 2^{n-t} times in a period of the sequence.

A direct consequence:

The sequence $\mathbf{a}$ has an *ideal ℓ-tuple distribution*, then $\mathbf{b}$ has t-tuple distribution for all $1 \leq t \leq \ell$.

m-sequences

An m-sequence possesses all the above randomness properties: 0-1, balance, run distribution and ideal n-tuple distribution.

Randomness properties required in a key stream generator in stream ciphers

Practical randomness criteria

1. Long period
2. Statistic properties
   (a) Balanced property
   (b) Run property R2
   (c) Ideal k-tuple distribution
3. Correlation
   (a) 2-level autocorrelation
   (b) Low crosscorrelation: $|C_{\mathbf{a},\mathbf{b}}(\tau)| \leq c\sqrt{N}$ where c is constant.
4. Large linear span:

$$t(N/2) \leq LS(\mathbf{a}) \leq N, \text{where } t \text{ is constant with } 0 < t \leq 1$$
$$t/2 < \rho \leq 1, \rho = LS(\mathbf{a})/N$$

Randomness properties required in a key stream generator in stream ciphers (cont.)

Indistinguishibility

- The probability that it can be distinguished from uniformly distributed sequences using any probabilistic polynomial time algorithms is negligible.

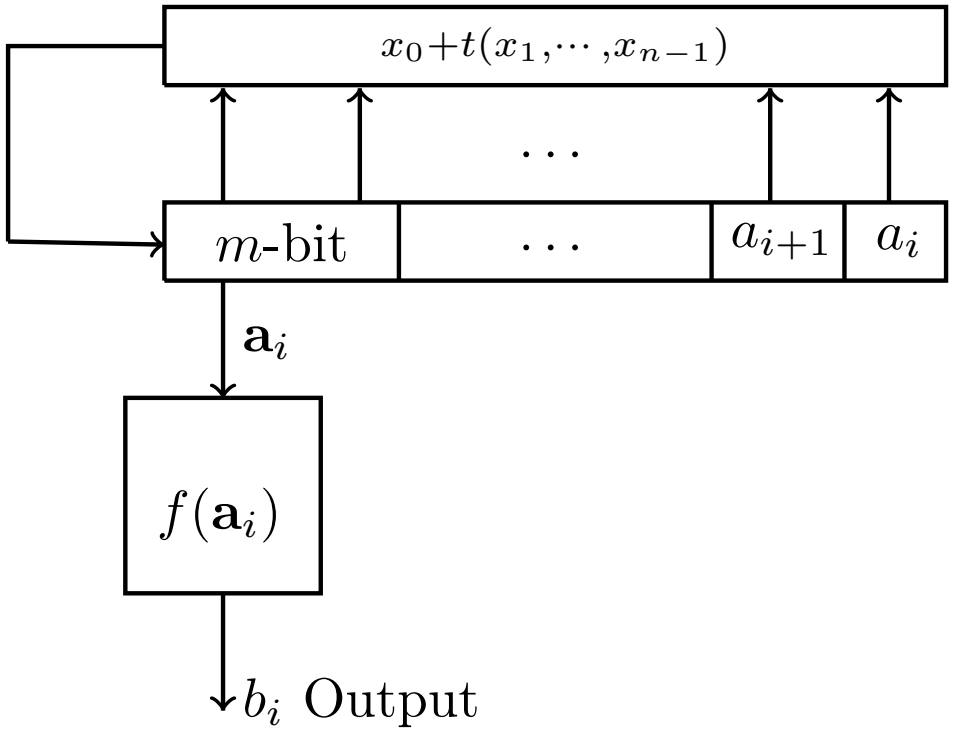
Why a de Bruijn sequence generator cannot be used directly?

- Randomness of de Bruijn sequences:
  - period 2^n
  - balanced
  - linear complexity at least $2^{n-1} + n$
  - ideal t-tuple distribution $1 \leq t \leq n$
- A de Bruijn sequence **cannot** be directly used as a PRNG or a keystream generator directly:
  - it outputs an internal state if the number of consecutive key stream bits is larger than b or the internal states can be recovered easily.
- This can be prevented by applying a **filtering function**!
- Challenge:

Can we still preserve ideal tuple distributions for filtering sequences?

Filtering de Bruijn sequences (FDBG)

- f is a boolean function in m variables. It takes consecutive m-bits from the last m stages of the de Bruijn generator.



- **m-tuple input to f**: $\mathbf{a}_i = (a_{i+n-m}, \cdots, a_{i+n-1}) \in \mathbb{F}_{2^m}, i \geq 0.$
- **Filtering sequence**: $b_i = f(\mathbf{a}_i), i = 0, 1, \cdots.$

How easy it is that the uniformity of tuple distribution of DB sequences can be preserved?

Proposition. The output sequence of FDBG generator ideal $(n-m+1)$-tuple distribution if x_0 or x_{m-1} in f is independent of the other variables:

$$\begin{aligned} f(x_0, \cdots, x_{m-1}) &= x_0 + g(x_1, \cdots, x_{m-1}) \text{ or} \\ f(x_0, \cdots, x_{m-1}) &= x_{m-1} + g(x_0, \cdots, x_{m-2}) \end{aligned}$$

where g is any Boolean function in $(m-1)$ variables.

- The above result can be obtained using the result from Golic (1998) and Canteau (2005).
- Since g can be any boolean function in $(m-1)$ variables, we should be able to obtain some g such that f possessing other strong cryptographic properties!

Approaches

- Looking for a family of filtering functions with good cryptographic properties, but preserve ideal tuple distribution property!
- We consider the set of functions whose evaluations have ideal two-level autocorrelation functions.
- Those functions are also called **orthogonal functions**.
- Currently all the known constructions of such functions are collected in Golomb-Gong's book, entitled as Signal Design with Good Correlation (2005), which will be listed in the next few slides.

Known constructions of binary 2-level autocorrelation sequences

Before 1997

- Number theory based: the Quadratic residue sequences (1932) of period p, p is prime; Hall's sextic residue sequences (1956) when $p = 4a^2 + 27$; and twin prime sequences.
- All $n \geq 2$, MLFSR-sequences = PN-sequences = m-sequences with period $N = 2^n - 1$ (Singer difference set (1932) and Golomb (1954)).
- If $n \geq 6$, n composite, GMW sequences of period $2^n - 1$ (1962, 1984).
- Exhaustive search done at $n = 7, 8, 9$, and 10.

Known constructions of binary 2-level autocorrelation sequences (cont.)

After 1997

- Conjectured Sequences (No-Golomb-Gong-Lee-Gaal 1998): 3-term sequences (also Gong-Gaal-Golomb, 1997) 5-term sequences, and Welch-Gong transformation sequences (Golomb-Gong-Gaal, 1998, No-Chung-Yun, another representation, 1999; proved by Dillon for odd case in 1999, and Dillon and Dobbertin for all cases (2004))
- Hyper-oval Construction: Segre case and Glynn I and II cases (Maschietti, 1998)
- Kasami Power Function Construction (Dobbertin 1998 for construction, Dillon and Dobbertin for proof, 2004, including 3-term and 5-term sequences as subclasses)
- Subfield Constructions

Known Constructions for p-ary Sequences of Period $p^n - 1$ with 2-level Autocorrelation for odd prime p

- m-sequences (Zieler, 1959)
- GMW Sequences (Gold-Miller-Welch, 1961)
- HG Sequences (Helleseth-Gong, 2002; the ternary case, Helleseth-Kumar-Martinsen, 2001; Dillon independently obtained HG sequences)
- Ternary case: Lin conjecture (Hu-Gao-Gong-Helleseth (2014), and Arasu-Dillon-Player (2014) together with the other conjectures by Ludkovski-Gong (2001))
- Subfield Constructions (No, 2002, for some special case)

Sequences, polynomial functions and boolean functions

- $\mathbb{F}_{2^m}$ denotes a finite field defined by an irreducible polynomial $p(x)$. In this talk, we assume that $p(x)$ is primitive and with $p(\alpha) = 0$.
- $Tr(x) = x + x^2 + \cdots + x^{2^{m-1}}$ denotes the trace function from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$.
- Let $\{a_i\}$ be a binary sequence of of period $2^m - 1$, then there exists a polynomial function $f(x) : \mathbb{F}_{2^m} \to \mathbb{F}_2$ such that $a_i = f(\alpha^i), i = 0, 1, \cdots$, and $f(x)$ is called a **trace representation** of the sequence where

$$f(x) = \sum_{i=1}^{r} \mathrm{Tr}_1^{m_i}(x^{t_i})$$

  where t_i is a coset leader modulo $2^{m_i} - 1$ and $m_i | m$.
- For $x = x_0 + x_1\alpha + \cdots + x_{m-1}\alpha^{m-1}, x_i \in \mathbb{F}_2$, $f(x)$ becomes a **boolean function** in m variables $x_0, \cdots, x_{m-1}$.

Example

m-sequences of periods 7 and 15, resp.:

$\mathbb{F}_{2^3} : \alpha^3 + \alpha + 1 = 0$:

$$\{a_i\} = 1001011 \quad \leftrightarrow Tr(x) = x_0$$

$$\{b_i\} = 1110100 \quad \leftrightarrow Tr(x^3) = x_0 + x_1 + x_2 + x_1x_2$$

$\mathbb{F}_{2^4} : \alpha^4 + \alpha + 1 = 0$:

$$\{c_i\} = 000100110101111 \quad \leftrightarrow f_c(x) = Tr(x) = x_3$$

$$\{b_i\} = 011110101100100 \quad \leftrightarrow f_d(x) = Tr(x^7)$$
$$= x_1 + x_2 + x_3 + x_0x_3 + x_1x_3 + x_2x_3 + x_1x_2x_3$$

Generalization of the Welch-Gong (WG) Transformation

Definition

Let p be a prime and $g(x)$ be a mapping from $\mathbb{F}_{p^m}$ to $\mathbb{F}_p$ which can be written as $g(x) = Tr(t(x))$ where $t(x)$ is a mapping of $\mathbb{F}_{p^m}$. The WG transformation of g is defined as

$$WG_g(x) = \mathbf{Tr}\big(t(x-a) - b\big) \textbf{ for fixed } a, b \in \mathbb{F}_{p^m}$$

- When $p = 2$, and $a = b = 1$, it becomes the original WG transformation.

E.g. For $m = 5$,

$g(x) = Tr(x^7) \leftrightarrow$ **an m-sequence of period 31**

$WG_g(x) = Tr((x+1)^7 + 1) = Tr(x + x^5 + x^7)$
$\leftrightarrow$ **a quadratic residue sequence of period 31**

Conjecture on WG Transformations

Conjecture (Mandal *et al.*, 2016)

For an odd m, let $WG(x)$ be the original WG transformation, then there exists a decimation $d = 2^J - 1$ such that the output of FDBG has up to ideal $(n - m + 1)$-tuple distributions where n is the length of the NLFSR when $f(x) = WG(x^d)$.

In this talk, we present a **more general case**. Let KPF be the set consisting of all Kasami power functions. We investigate $g \in KPF$ and the WG transformation of g as a filtering function in FDBG.

Kasami Power Function Construction

KPF (Dillon-Dobbertin, 2004)

Let m be a positive integer and k and k' be such that $kk' \equiv 1 \bmod m$, $k < m$ and $\gcd(k, m) = 1$. Let $R_{k'}(x)$ be a permutation polynomial over $\mathbb{F}_{2^m}$, which is recursively defined as [1]

$$R_{k'}(x) = \sum_{i=1}^{k'} A_i(x) + V_{k'}(x)$$

where $A_i(x)$ and $V_i(x)$ are defined below:

$$A_1(x) = x,\ A_2(x) = x^{2^k+1}$$
$$A_{i+2} = x^{2^{(i+1)k}} A_{i+1}(x) + x^{2^{(i+1)k}-2^{ik}} A_i(x), i \geq 1 \text{ and}$$
$$V_0(x) = 0,\ V_2(x) = x^{2^k-1}$$
$$V_{i+2} = x^{2^{(i+1)k}} V_{i+1}(x) + x^{2^{(i+1)k}-2^{ik}} V_i(x), i \geq 1.$$

Then $\mathbf{Tr}(R_{k'}(x))$ gives a 2-level autocorrelation sequence.

[1] J.F. Dillon and H. Dobbertin. New Cyclic Difference Sets with Singer Parameters, Finite Fields and Their Applications, vol. 10, Issue 3, pp. 342 – 389, Elsevier, 2004

3-term and 5-term Sequences from the KPF

The 3-term and 5-term sequences can be obtained from the KPF by setting $k' = 2$ and 3, respectively.

- **3-term sequences:** For $k' = 2$, $2k \equiv 1 \bmod m$, then

$$T3(x) = \mathrm{Tr}(R_2(x)) \text{ where}$$

$$R_2(x) = x + x^{2^k+1} + x^{2^k-1}$$

- **5-term sequences:** For $k' = 3$, $3k \equiv 1 \bmod m$, then

$$T5(x) = \mathrm{Tr}(R_3(x)) \text{ where}$$

$$R_3(x) = x + x^{2^k+1} + x^{2^{2k}+2^k+1} + x^{2^{2k}-2^k+1} + x^{2^{2k}+2^k-1}$$

The Welch-Gong (WG) Transformation Sequences

WG transformation sequences:

- The Welch-Gong (WG) transformation from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$ is given by

$$WG_{T5}(x) = \mathbf{Tr}(R_3(x+1)+1), x \in \mathbb{F}_{2^m}.$$

- A WG transformation sequence has 2-level autocorrelation, so does its decimation:

$$WG_{T5}(x^d) = \mathbf{Tr}(R_3(x^d+1)+1), \gcd(d, 2^m-1) = 1$$

E.g. $m = 7$,

$$T5(x) = Tr(R_3(x)) = Tr(x + x^{33} + x^{39} + x^{41} + x^{104}),$$

$$\{T5(\alpha^i)\} - \textbf{5-term sequence}$$

$$WG_{T5}(x) = Tr(R_3(x+1)+1) = Tr(x + x^3 + x^7 + x^{19} + x^{29}),$$

$$\{WG_{T5}(\alpha^i)\} - \textbf{WG sequence}$$

WG sequence with desimatio 7:

$$WG_{T5}(x^7) = Tr(R_3(x^7+1)+1) = Tr(x^3 + x^7 + x^{11} + x^{19} + x^{21})$$

They all have 2-level autocorrelation (decimation preserves the 2-level autocorrelation).

Does the WG transform preserve 2-level autocorrelation?

Answer.

- It is a **NO**!
- For f giving a 2-level autocorrelation sequence, the WG transform of f produces a 2-level autocorrelation sequence **only when** $f = T5$, at least experimentally verified. Those sequences are originally the so-called WG sequences.

New Results on WG Transform and Ideal Tuple Distribution

New Property: Invariant under the WG Transform

Definition

We say a function $f(x) : \mathbb{F}_{2^m} \to \mathbb{F}_2$ is ***invariant*** under the WG transform if the WG transform of $f(x)$ is equal to the function itself for m odd or its complement for m even, i.e.,

$$WG_f(x) = \begin{cases} f(x) & m \textbf{ odd} \\ f(x)+1 & m \textbf{ even} \end{cases}$$

E.g Let $m = 5$

$$f(x) = Tr(x^3 + x^{15} + x^{21})$$

$$\begin{aligned} Tr((x+1)^3) &= Tr(x^3 + 1) \\ Tr((x+1)^{15}) &= Tr(x^{15} + x^{12} + x^9 + 1) \\ Tr((x+1)^{21}) &= Tr(x^{21} + x^{20} + x^{17} + 1) \end{aligned}$$

$$WG_f(x) = Tr((x+1)^3 + (x+1)^{15} + (x+1)^{21} + 1) = f(x)$$

Thus f is invariant under the WG transform.

WG invariants

Theorem 1.

For $m > 6$, the function $WG_f(x^d)$ is invariant under the WG transform for $f(x) \in \{T3(x), T5(x)\}$ and $d = 2^{m-k+1} - 1$.

- Theorem 1 is proved using Hadamard transform of $WG_f(x^d)$.

E.g. For $m = 7$,

- 3-term case:

$$m = 7, 2k \equiv 1 \bmod 7 \Longrightarrow k = 4, d = 2^4 - 1 = 15$$

$$T3(x) = Tr(R_2(x)), R_2(x) = x + x^{17} + x^{15}$$
$$WG_{T3}(x) = Tr(R_2(x+1) + 1) = Tr(x + x^3 + x^{11} + x^{13} + x^{15})$$

$$WG_{T3}(x^{15}) = Tr(R_2(x^{15}+1)+1) = Tr(x^9+x^{11}+x^{15}+x^{19}+x^{45})$$

- 5-term case:

$$m = 7, 3k \equiv 1 \bmod 7 \Longrightarrow k = 5, d = 2^3 - 1 = 7$$
$T5(x), WG_{T5}(x), WG_{T5}(x^7)$, given previously, reproduced below

$$WG_{T5}(x^7) = Tr(R_3(x^7 + 1) + 1) = Tr(x^3 + x^7 + x^{11} + x^{19} + x^{21})$$

- f is invariant under WG transform

$$WG_f(x) = f(x), f \in \{WG_{T3}(x^{15}), WG_{T5}(x^7)\}$$

Hadamard Transform and Its Inverse

- f, a boolean function in m variables.
- **Hadamard transform**

$$\hat{f}(\lambda) = \sum_{x \in \mathbb{F}_{2^m}} (-1)^{\mathbf{Tr}(\lambda x)+f(x)}, \forall \lambda \in \mathbb{F}_{2^m},$$

- **The inverse Hadamard transform**

$$(-1)^{f(x)} = \frac{1}{2^m} \sum_{\lambda \in \mathbb{F}_{2^m}} (-1)^{\mathbf{Tr}(\lambda x)} \hat{f}(\lambda),$$

Relation Between Hadamard Transforms of f and WG_f

- $$\widehat{WG}_f(\lambda) = (-1)^{\mathbf{Tr}(\lambda+1)}\hat{f}(\lambda), \forall \lambda \in \mathbb{F}_{2^m}.$$

- Consequently, the magnitude of f and its WG transform are the same.
- So, the following is true.

> If $\hat{f}(\lambda) \neq 0$ implies that $\mathrm{Tr}(\lambda) = 1$, then
> $$\widehat{WG}_f(\lambda) = \hat{f}(\lambda), \forall \lambda \in \mathbb{F}_{2^m} \Longrightarrow WG_f(x) = f(x)$$

Known Results on Hadamard Transform of WG Transform

Hadamard transform of WG (Dillon, 1999) [2]

For $3k \equiv 1 \bmod m$ and m odd,

$$\widehat{WG}_{T5}(\lambda) = \begin{cases} 0 & \text{if } \mathrm{Tr}(\lambda^{d^{-1}}) = 0 \\ \pm 2^{\frac{m+1}{2}} & \text{if } \mathrm{Tr}(\lambda^{d^{-1}}) = 1 \end{cases}$$

where $d = 2^{2k} - 2^k + 1$, i.e., $WG_{T5}(x)$ has three valued Hadamard transform (called preferred three values or almost bent).

[2] J.F. Dillon. Multiplicative Difference Sets via Additive Characters, Designs, Codes and Cryptography, vol. 17, Issue 1, pp. 225 – 235, 1999.

How about decimation of $WG_{T5}(x)$?

Unique decimation of WG for preferred Hadamard transform (Yu-Gong, 2006)

Let $d_1 = \frac{2^{2k}-2^k+1}{2^k+1}$ and $f(x) = WG_{T5}(x^{d_1})$. Then the Hadamard transform of the WG transform [3] with decimation d_1 is three valued:

$$\widehat{f}(\lambda) \in \{0, \pm 2^{\frac{m+1}{2}}\}$$

and d_1 is the only decimation such that $WG_{T5}(x^{d_1})$ has the preferred three valued Hadamard transform.

[3] N.Y. Yu and G. Gong. Crosscorrelation Properties of Binary Sequences with Ideal Two-level Autocorrelation, In SETA'06, LNCS 4086, pp. 104 – 118, Springer, Berlin, Heidelberg, 2006.

New Result on the Hadamard Transform

Lemma (Mandal and Gong, 2017)

For $f(x) = WG_{T5}(x^{d_1})$,

$$\widehat{f}(\lambda) = \begin{cases} 0 & \text{if } Tr(\lambda) = 0 \\ \pm 2^{\frac{m+1}{2}} & \text{if } Tr(\lambda) = 1. \end{cases}$$

The proof can be obtained using a similar approach as Dillon did.

Connection between WG decimations

- We determine that the decimation d of Mandal et al.'s Conjecture is $d = 2^{m-k+1} - 1$, which is in Theorem 1.
- It turns out that d and d_1 are in the same coset.
- Thus, for $f(x) = WG_{T5}(x^d)$, then the Hadamard transform of f and the Hadamard transform of the WG transform of f are equal.
- So, the result in Theorem 1 for WG_{T5} is established.

The WG Transform of Three-term Functions

Known Results on Three-term Sequences $m = 2k - 1$.

- Hadamard transform of $T3(x^{2^k+1})$ is three valued: $\{0, \pm 2^{\frac{m+1}{2}}\}$ (Dillon-Dobbertin, 2004).
- Chang-Gaal-Golomb-Gong-Helleseth-Kumar[4] proved the code corresponding to $T3(x^{2^k+1})$ have 5 weights: $\{0, \pm 2^{\frac{m+1}{2}}, \pm 2^{\frac{m+3}{2}}\}$.
- Yu and Gong (2006) found another decimation $d = 2^k - 1$, the Hadamard transform of $T3(x^d)$ is also at most 5-valued, i.e., $\{0, \pm 2^{\frac{m+1}{2}}, \pm 2^{\frac{m+3}{2}}\}$.

[4] A. Chang and P. Gaal and S.W. Golomb and G. Gong and T. Helleseth and P.V. Kumar. On a Conjectured Ideal Autocorrelation Sequence and a Related Triple-error Correcting Cyclic Code, IEEE Transactions on Information Theory, vol. 46, No. 2, pp. 680 – 687, 2000.

New Results on the Hadamard Transform of the WG Transform of Three-term Functions and Invariants

Property 1 (Mandal and Gong, 2017)

For $m = 2k - 1$ and $d = 2^{m-k+1} - 1$, the Hadamard transforms of $WG_{T3}(x)$ and $f(x) = WG_{T3}(x^d)$ are 5-valued, i.e.,

$$\widehat{WG_{T3}}(\lambda), \hat{f}(\lambda) \in \{0, \pm 2^{\frac{m+1}{2}}, \pm 2^{\frac{m+3}{2}}\}.$$

Using this result, the second part of Theorem 1 for the invariant of the WG transform of $T3$ with decimation d can be established, i.e., $WG_f(x) = f(x)$.

Remark. $WG_{T3}(x)$ does not have 2-level autocorrelation, nor does $WG_{T3}(x^d)$. But the latter is invariant under the WG transform.

Boolean Representation and Invariants of the WG Transform

Boolean Representation

Let f be a polynomial function from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$. Then f has the following representation

$$f(x) = x_0[f(z) + WG_f(z) + Tr(1)] + h(z), x = x_0 + z, \textbf{where}$$

$$z = x_1\alpha + \cdots + x_{m-1}\alpha^{m-1}, x_i \in \mathbb{F}_2.$$

where $h(z)$ only depends on $x_1, \cdots, x_{m-1}$.

Property 2

The boolean representation of $f(x)$ has the following form

$$f(x) = x_0 + h(z)$$

if and only if $f(x)$ is invariant under the WG transform, i.e.,

$$WG_f(x) = f(x).$$

Ideal Tuple Distribution of WG Transformations in FDBG

Theorem 2 (Mandal and Gong, 2017)

For $k'k \equiv 1 \bmod m$ with $k' = 2, 3$ and m odd, the WG transform of $g \in \{T3, T5\}$ with decimation $d = (2^{m-k+1} - 1)$, denoted by $f(x) = WG_g(x^d)$, can be written as

$$f(x) = x_0 + h(x_1, x_2, \cdots, x_{m-1})$$

where $x = x_0 + x_1\alpha + \cdots + x_{m-1}\alpha^{m-1} \in \mathbb{F}_{2^m}$ and $h(x_1, \cdots, x_{m-1})$ is independent of x_0.

Ideal Tuple Distribution of WG Transformations in FDBG (cont.)

Theorem 3 (Mandal and Gong, 2017)

With the above notion, the filtering de Bruijn sequence $\mathbf{b}$ with $WG_h(x^d), h \in \{T3, T5\}$ as a filtering function in the FDBG has an ideal t-tuple distribution where $t = (n - m + 1)$ and n is the length of the NLFSR generating a de Bruijn sequence.

The case $h = T5$ asserts the conjecture.

What is t-tuple distribution for the other KPF functions in FDBG?

Experimental Results on KPFs with $k' = 4, 5$

- We use $R_{k'}(x^d)$ and $WG_{R_{k'}}(x^d)$ including all decimations d as filtering functions in the filtering de Bruijn generators and compute the tuple distribution.

- For $k' = 4$, $WG_{R_4}(x^d) = \mathrm{Tr}(R_4(x^d + 1) + 1), x \in \mathbb{F}_{2^m}$.

Input size m	NLFSR length n	t-tuple dist.
9	17	2
11	17	2
13	17	2

- For $k' = 5$, $WG_{R_5}(x^d) = \mathrm{Tr}(R_5(x^d + 1) + 1), x \in \mathbb{F}_{2^m}$.

Input size m	NLFSR length n	t-tuple dist.
9	17	2
11	17	2
13	17	2

How about even case of m?

Experimental Results on KPFs with $k' = 4, 5$ for m even: The experiments results show that for m even, no KPF function f used in FDBG can produces ideal t-tuple distribution for $t > 2$.

- We tested $R_{k'}(x^d)$ and $WG_{R_{k'}}(x^d)$ including all decimations d as filtering functions in the filtering de Bruijn generators and compute the ideal tuple distribution.
- When $k' = 3$ and m is even

Input size m	NLFSR length n	t-tuple dist.
8	9 – 17	2
10	11 – 17	2
14	15 – 17	2

- When $k' = 5$ and m is even

Input size m	NLFSR length n	t-tuple dist.
8	9 – 17	2
12	13 – 17	2
14	15 – 17	2

Concluding Remarks and Open Problems

- In this talk, we introduced the invariant under the WG transform of a Boolean function.
- We showed that $WG_{T3}(x^d)$ and $WG_{T5}(x^d)$ are **invariant** under the WG transform where $d = 2^{m-k+1} - 1$ for m odd.
- Using this property, we obtain that the FDBG has **ideal t-tuple distribution for $t = n - m + 1$** when each of the above two functions used as a filtering function in the FDBG, which also asserts the conjecture.
- For **all orthogonal functions of the KPF class**, the experiments show that there are no filtering functions in FDBG which can produce the ideal t-tuple distribution for $t > 2$ for m even as well as for m odd except for the two cases above.
- We also obtain two new classes of functions, i.e., $WG_{T3}(x^r), r \in \{1, d = 2^{m-k+1} - 1\}$ have **5-valued Hadamard transform** for m odd.

Hadamard Transform and Invariant under WG

- It is rather surprised that $d = 2^{m-k+1} - 1$ is **unique** decimation number found by Yu and Gong in their 2006 paper such that $WG_{T5}(x^d)$ has 3-valued Hadamard transform (m odd) and $T3(x^d)$ has 5-valued Hadamard transform. (It has a slightly different form in that paper.)

- It seems that the invariants of the WG transform have some relation with their Hadamard transform.

Differentials and Invariant under WG

- **Ideal t-tuple distribution property** in FDBG solely replies on the invariant property of the filtering functions under the WG transform.
- From **differential cryptanalysis**, f is invariant under WG means that the differential f

$$D_a(f) = f(x) + f(x+a)$$

  being constant 1 when $a = 1$.
- Nevertheless, we do not know how to use this property to attack FDBG when it is used as a key stream generator in **stream ciphers**.
- So, it may exist some **trade-offs** between the uniformity of the tuple distribution of FDBG and differential property.

Some Open Problems

The results in the experiments are true for general m. Precisely, we may form the following two conjectures. Let KPF be the set consisting of all KPFs.

Conjecture

For m odd or even,

$$h(x^d), WG_h(x^d) = h(x^d + 1) + Tr(1), \forall h \in KPF$$

with all decimations except for those two cases in Theorem 1, are not invariant under WG.

Some Open Problems (cont.)

Classification of invariants under WG transform

More generally, given a polynomial function $f(x) : \mathbb{F}_{2^m} \to \mathbb{F}_2$, how to efficiently determine whether the function is invariant under the WG transform, i.e., $f(x) = WG_f(x)$ for m odd and $f(x) = WG_f(x) + 1$ for m even.

Thank you!

Happy Birthday to Tor!

Communication Security (ComSec) Lab
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, ON, N2L 3G1, CANADA
www.comsec.uwaterloo.ca/~ggong