## Perturbations of binary de Bruijn sequences and spectral theory of graphs

### Martianus Frederic Ezerman, Adamas Aqsa Fahreza NTU, Singapore <u>Janusz Szmidt</u> Military Communication Institute, Poland

The International Workshop on Boolean Functions and their Applications  $$\operatorname{BFA}$$  2019

June 2019

### The Feedback Shift Registers - FSRs

Let 𝔽<sub>2</sub> be the binary field and 𝔽<sub>2</sub><sup>n</sup> the *n*-dimensional vector space over 𝔽<sub>2</sub>. Let us consider a mapping

$$\mathfrak{F}:\mathbb{F}_2^n\to\mathbb{F}_2^n$$

$$\mathfrak{F}(x_0,\ldots,x_{n-1}) = (x_1,x_2,\ldots,x_{n-1},f(x_0,\ldots,x_{n-1}))$$
 (1)

where f is a Boolean function of n variables of the form

$$f(x_0,\ldots,x_{n-1}) = x_0 + F(x_1,\ldots,x_{n-1}),$$
 (2)

and F is a Boolean function of n-1 variables.

- ▶ The condition (2) defines a nonsingular *FSR* of order *n*.
- ► A nonsingular register decomposes the space 𝔽<sub>2</sub><sup>n</sup> into a finite number of disjoint cycles.

## de Bruijn Sequences

- If there is only one cycle (of length 2<sup>n</sup>), then we have a de Bruijn sequence.
- The number of cyclically non-equivalent de Bruijn sequences of order n is (published by de Bruijn, 1946)

$$B_n = 2^{2^{n-1} - n} (3)$$

- In fact, these sequences were discovered by Fench mathematician C. Flye Sainte-Marie in 1984 and he proved formula (3).
- Consider a de Bruijn sequence s = (s<sub>0</sub>, s<sub>1</sub>,...) with given *n*-initial elements (s<sub>0</sub>,..., s<sub>n-1</sub>). The next elements, for i ≥ 0, are calculated from the formula

$$s_{i+n} = f(s_i, s_{i+1}, \dots, s_{i+n-1}) = s_i + F(s_{i+1}, \dots, s_{i+n-1})$$
 (4)

for some Boolean function F.

### Perturbation of Boolean Functions

- ▶ Let S(n) be the set of functions  $\mathbb{F}_2^{n-1} \to \mathbb{F}_2$ , which generate de Bruijn sequences of order *n*. For a function  $F : \mathbb{F}_2^{n-1} \to \mathbb{F}_2$ , we define the set S(F; k) of functions  $g \in S(n)$  such that the weight of the function F + g equals *k*. It means that the number of inputs for which the functions *F* and *g* are different equals *k*.
- We introduce the notation N(F; k) = |S(F; k)|
- and dwfine the counting function

$$G(F; y) = \sum_{k} N(F; k) y^{k}.$$
 (5)

### The Counting Formula

- D. Coppersmith, R. C. Rhoades, J. M. Vanderkam. Counting de Bruijn sequences as perturbations of linear recursions (arXiv e-prints, May 2017) have proved the formula:
- ► Let ℓ be a linear function, then they have proved Theorem

$$G(\ell; y) = \sum_{k} N(\ell; k) y^{k} = 2^{-n} \prod_{c \in C(\ell)} p_{d(c)},$$
(6)

where  $C(\ell)$  denotes the set of all cycles generated by the feedback function  $\ell(x_0, x_1, \ldots, x_{n-1})$ , for any  $c \in C(\ell)$  let d(c) be the number of ones in the cycle c and

$$p_k(y) = (1+y)^k - (1-y)^k$$
 for  $k > 0$  and  $p_0(y) = 1$ .

The proof of Coppersmith et al. formula (6)

Let G<sub>n</sub> be the de Bruijn graph of order n. It is a 2-in and 2-out directed graph with 2<sup>n</sup> vertices corresponding to elements of ℝ<sup>n</sup><sub>2</sub> and an edge

$$x_0x_1\ldots x_{n-1} \longrightarrow x_1\ldots x_{n-1}x_n$$

for all choices of  $x_0, x_1, \ldots x_{n-1}, x_n \in \mathbb{F}_2$ .

- Let F : ℝ<sub>2</sub><sup>n-1</sup> → ℝ<sub>2</sub> be a Boolean function. In the graph G<sub>n</sub> label the edge x<sub>0</sub>x<sub>1</sub>...x<sub>n-1</sub> → x<sub>1</sub>...x<sub>n-1</sub>x<sub>n</sub> by 1 if x<sub>n</sub> = x<sub>0</sub> + F(x<sub>1</sub>,...,x<sub>n-1</sub>) and by y otherwise. Denote the weighted graph by G<sub>n,F</sub>.
- Denote the weighted adjacency matrix by W<sub>n,F</sub>. Hence W<sub>n,F</sub> is the 2<sup>n</sup> × 2<sup>n</sup> matrix with 1 in the row and column corresponding to x<sub>0</sub>x<sub>1</sub>...x<sub>n-1</sub> → x<sub>1</sub>...x<sub>n-1</sub>x<sub>n</sub> where x<sub>n</sub> = x<sub>0</sub> + F(x<sub>1</sub>,...,x<sub>n-1</sub>) and y otherwise.

## The proof of Coppersmith et al. formula (6)

If we view x<sub>0</sub>...x<sub>n-1</sub> as the binary representation of some integer *i* then W<sub>n,F</sub>(*i*,*j*) = 1 if and only if x<sub>1</sub>,..., x<sub>n1</sub>x<sub>n</sub> is the binary representation of *j*, W<sub>n,F</sub>(*i*, *k*) = *y* if and only if x<sub>1</sub>,..., x<sub>n-1</sub>x<sub>n</sub> is the binary representation of *k*.

#### Lemma

Let M be a  $p \times p$  matrix such that the sum of the entries in every row and column is 0. Let  $M_0$  be the matrix obtained from M by removing the first row and first column. Then the coefficient of z in the characteristic polynomial det(M - zI)(with I the identity matrix) of M is equal to  $-p \cdot det(M_0)$ .

## The proof of Coppersmith et al. formula (6)

They have proven the following Proposition which is valid for arbitrary Boolean function  $F : \mathbb{F}_2^{n-1} \to \mathbb{F}_2$ , hence for any non-singular *FSR*  $f = x_0 + F$ 

#### Proposition

 $(1+y)^{2^{n-1}-1} \cdot G(F; y)$  is equal to the determinant of  $(1+y)I_n - W_{n,F}$  after deleting the first row and column.

### The Fryers Formula

Let a linear function l : F<sub>2</sub><sup>n-1</sup> → F<sub>2</sub> generate the *m*-sequence of the period 2<sup>n</sup> − 1. Then the right side of (6) has the form

$$G(\ell; y) = \frac{1}{2^n} \left( (1+y)^{2^{n-1}} - (1-y)^{2^{n-1}} \right)$$
(7)

This formula was attributed to Michael Fryers.

- ► The coefficients N(l; k) can be calculated by expanding the powers in G(l; y).
- The coefficients  $N(\ell, k)$  have the following interpretation.
- N(ℓ; 1) = 1. Hence, from an *m*-sequence we get one de Bruijn sequence by adding the cycle of the zero state, corresponding to one change in the truth table of the function ℓ.
- N(ℓ; 2) = 0. The truth table of ℓ is changed in two places.
   One change adds the zero cycle and the second cuts the full cycle into two cycles. No new de Bruijn sequence obtained.
- In general, N(ℓ; k) = 0 for all even k since an even number of changes in the truth table of ℓ always lead to disjoint cycles.

### The Fryers Formula - A Combinatorial View

- ► There is in fact an interesting combinatorial view on the non-vanishing (> 0) coefficients of the polynomial G(ℓ, y).
- Sequence A281123 in OEIS gives the formula for the positive coefficients of the polynomial

$$G(\ell; y) = q(n-1, y) = \frac{(1+y)^{2^{n-1}} - (1-y)^{2^{n-1}}}{2^n}$$

▶ Hence, for odd  $1 \le k \le 2^{n-1} - 1$ , the formula for  $N(\ell, k)$  is

$$N(\ell,k) = \frac{1}{2^{n-1}} \begin{pmatrix} 2^{n-1} \\ k \end{pmatrix} \text{ for } n \ge 2.$$
(8)

 More details and related integer sequences (such as Pascal Triangle) can be found in https://oeis.org/A281123

### The Fryers Formula - Corrolaries - 2

▶ The Helleseth and Kløve formula (1991) follows from (8)

$$N(\ell;3) = \frac{(2^{n-1}-1)(2^{n-1}-2)}{3!}$$

for the number of cross-join pairs for an *m*-sequence.

The higher Helleseth and Kløve formula

$$N(\ell; k = 2j + 1 \ge 5) = \frac{1}{k!} \prod_{i=1}^{k-1} (2^{n-1} - i)$$

gives the number of new de Bruijn sequences obtained after the *j*-th application of cross-join method: (a) start from an *m*-sequence, (b) add 0 to obtain a de Bruijn sequence, (c) find all of its cross-join pairs, (d) use them to construct new de Bruijn sequences, (e) for each resulting sequence, repeat (c) and (d) j - 1 times.

### The Fryers Formula - Corrolaries - 3

Using (8) we easily obtain the number of all cyclically non-equivalent de Bruijn sequences of order n:

$$G(\ell, 1) = \sum_{k=1}^{2^{n-1}-1} N(\ell, k) = \frac{1}{2^{n-1}} \underbrace{\sum_{k=1}^{2^{n-1}-1} \binom{2^{n-1}}{k}}_{:=\alpha} = 2^{2^{n-1}-n}$$
Ince  $\alpha = 2^{2^{n-1}-1}$  is the sum of the odd entries in row  $2^{n-1}$  or

since  $\alpha = 2^{2^{n-1}-1}$  is the sum of the odd entries in row  $2^{n-1}$  of the Pascal Triangle.

▶ For *n* = 4

$$\sum_{k=1}^{7} N(\ell, k) = 1 + 7 + 7 + 1 = 2^{4}.$$

For n = 5 and n = 6

$$\sum_{k=1}^{15} N(\ell, k) = 1 + 35 + 273 + 715 + 715 + 273 + 35 + 1 = 2^{11}.$$

▶ For *n* = 6

$$\sum_{k=1}^{31} N(\ell, k) = 1 + 155 + 6293 + 105183 + 876525 + 4032015$$

 $\begin{array}{l}+10855425+17678835+17678835+10855425+4032015\\\\+876525+105183+6293+155+1=67108864=2^{26}\end{array}$ 

#### The coefficients are symmetric.

### Starting from non-primitive linear recurrences

- Versions of Fryers' formula (7) can be derived from Coppersmith et al formula (6) for non-primitive linear recurrences. We need to that the cyclic decomposition of the generated sequence and the weights of cycles (the number of ones on the cycles).
- Perturbation of linear recurrences for the order n = 4

Recurrence	Fryers' Polynomial	# of de Bruijn Sequences
$x_0 + x_1$	$y + 7y^3 + 7y^5 + y^7$	(1,7,7,1)
$x_0 + x_3$	$y + 7y^3 + 7y^5 + y^7$	(1, 7, 7, 1)
<i>x</i> <sub>0</sub>	$12y^{5} + 4y^{7}$	(12, 4)
$x_0 + x_2$	$8y^3 + 8y^5$	(8,8)
$x_0 + x_1 + x_2$	$6y^3 + 8y^5 + 2y^7$	(6, 8, 2)
$x_0 + x_1 + x_3$	$12y^{5} + 4y^{7}$	(12, 4)
$x_0 + x_2 + x_3$	$6y^3 + 8y^5 + 2y^7$	(6,8,2)
$x_0 + x_1 + x_2 + x_3$	$8y^3 + 8y^5$	(8,8)

## Starting from non-primitive linear recurrences

- Pure Circulating Register (PCR)  $I = x_0$  for the order n = 5
- ► The Freyers' polynomial 576y<sup>7</sup> + 960y<sup>9</sup> + 448y<sup>11</sup> + 64y<sup>13</sup>
- ► The sum of coefficients 576 + 960 + 448 + 64 = 2048 = 2<sup>11</sup>
- 576 is the number of spanning trees for the adjacency graph of PCR.
- J. Mykkeltveit and J. Szmidt.

*On cross joining de Bruijn sequences*, Contemporary Mathematics, vol. 632, pp. 333-344 (2015).

#### Theorem

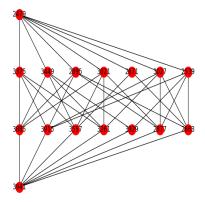
Let  $(u_t)$ ,  $(v_t)$  be two de Bruijn sequences of order n. Then  $(v_t)$  can be obtained from  $(u_t)$  by repeated applications of the cross-join operation.

### Starting from non-linear recurrences

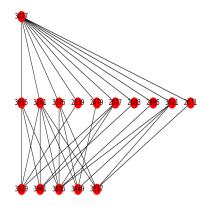
- When we start from non-singular and non-linear recurrence we can ganarate all de Bruijn sequences. First applying the joining of cycles and then the cross-join method.
- We consider a situation when we start from modified de Bruijn sequence (or corresponding non-linear recurrence).
- We first add the zero cycle which corresponds to term y in the Fryers polynomial and then apply the cross-join method until we generate all de Bruijn sequences.
- We have not a compact form of the right hand side of formula
   (6) but we can calculate the Fryers calculate

Two Patterns for n = 4

Symmetric (1, 7, 7, 1)



Asymmetric (1, 10, 5, 0)



### First 30 of Exactly 60 Distribution Patterns for Order 5

Coefficients	#	Coefficients	#
(1, 34, 276, 713, 713, 276, 34, 1)	192	(1, 34, 297, 804, 699, 202, 11, 0)	32
(1, 39, 310, 790, 677, 211, 20, 0)	112	(1,40,317,768,691,216,15,0)	32
(1, 37, 322, 770, 685, 217, 16, 0)	96	(1, 35, 294, 806, 701, 199, 12, 0)	32
(1, 35, 273, 715, 715, 273, 35, 1)	96	(1, 32, 278, 717, 709, 274, 36, 1)	32
(1, 32, 237, 640, 739, 352, 47, 0)	80	(1, 36, 315, 792, 671, 212, 21, 0)	32
(1, 45, 351, 743, 639, 235, 33, 1)	64	(1, 34, 235, 636, 743, 354, 45, 0)	32
(1, 36, 270, 717, 717, 270, 36, 1)	64	(1, 31, 260, 726, 737, 267, 26, 0)	32
(1, 37, 274, 706, 717, 281, 32, 0)	64	(1, 32, 261, 720, 739, 272, 23, 0)	32
(1, 33, 235, 639, 743, 351, 45, 1)	64	(1, 33, 275, 719, 711, 271, 37, 1)	32
(1, 33, 262, 714, 741, 277, 20, 0)	64	(1, 37, 271, 711, 719, 275, 33, 1)	32
(1, 33, 278, 714, 709, 277, 36, 0)	64	(1, 39, 316, 774, 689, 211, 18, 0)	32
(1, 47, 349, 739, 643, 237, 31, 1)	48	(1,48,301,672,675,304,47,0)	32
(1, 31, 237, 643, 739, 349, 47, 1)	48	(1, 49, 364, 834, 633, 157, 10, 0)	32
(1, 36, 274, 709, 717, 278, 32, 1)	32	(1, 47, 370, 830, 629, 163, 8, 0)	32
(1, 40, 341, 752, 659, 232, 23, 0)	32	(1,47,301,675,675,301,47,1)	32

### Last 30 of Exactly 60 Distribution Patterns for Order 5

Coefficients	#	Coefficients	#
(1, 35, 276, 710, 713, 279, 34, 0)	32	(1, 41, 352, 858, 649, 141, 6, 0)	16
(1, 35, 318, 790, 669, 215, 20, 0)	24	(1, 45, 372, 834, 625, 161, 10, 0)	16
(1, 43, 374, 838, 621, 159, 12, 0)	24	(1, 43, 352, 846, 665, 135, 6, 0)	16
(1, 45, 366, 850, 613, 161, 12, 0)	16	(1, 41, 314, 770, 693, 213, 16, 0)	16
(1, 42, 315, 764, 695, 218, 13, 0)	16	(1, 38, 319, 772, 687, 214, 17, 0)	16
(1, 39, 352, 870, 633, 147, 6, 0)	16	(1, 41, 382, 826, 629, 157, 12, 0)	16
(1, 45, 344, 858, 657, 137, 6, 0)	16	(1, 48, 349, 736, 643, 240, 31, 0)	16
(1, 47, 346, 846, 661, 147, 0, 0)	16	(1, 32, 269, 704, 739, 288, 15, 0)	8
(1, 34, 267, 700, 743, 290, 13, 0)	16	(1, 47, 366, 838, 629, 155, 12, 0)	8
(1, 37, 306, 770, 717, 217, 0, 0)	16	(1, 36, 265, 696, 747, 292, 11, 0)	8
(1, 51, 362, 830, 637, 159, 8, 0)	16	(1, 39, 382, 838, 613, 163, 12, 0)	8
(1, 35, 324, 774, 681, 215, 18, 0)	16	(1, 39, 330, 878, 677, 123, 0, 0)	8
(1, 33, 296, 810, 697, 197, 14, 0)	16	(1, 60, 401, 776, 603, 188, 19, 0)	8
(1, 62, 399, 772, 607, 190, 17, 0)	16	(1, 31, 282, 814, 725, 195, 0, 0)	8
(1, 35, 304, 782, 713, 207, 6, 0)	16	(1, 64, 397, 768, 611, 192, 15, 0)	8

### Patterns for n = 5

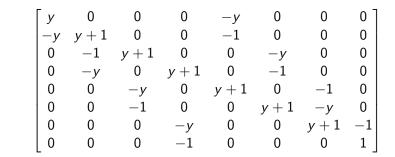
- For n = 5 there are exactly 60 perturbation patterns.
- The pattern in blue matches the Fryer's coefficients. Note that only 6 out of the 96 sequences correspond to the *m*-sequences.
- Other symmetric patterns are in bold.
- Each of the two pairs of patterns in red consists of reversals.
- Our investigation into various interesting patterns has only just begun.
- There are many open questions and directions.

The weighted adjacency matrices, n = 3,  $f = x_0 + x_2$ ,  $W_{f,3}$ 

An example, n = 3 and  $F = x_2$ , (The feedback function of the register (1) is  $f = x_0 + x_2$ ). Then the weighted matrix

$$W_{n,F} = egin{bmatrix} 1 & 0 & 0 & 0 & y & 0 & 0 & 0 \ y & 0 & 0 & 0 & 1 & 0 & 0 & 0 \ 0 & 1 & 0 & 0 & 0 & y & 0 & 0 & 0 \ 0 & y & 0 & 0 & 0 & 1 & 0 & 0 & 0 \ 0 & 0 & y & 0 & 0 & 0 & 1 & 0 & 0 \ 0 & 0 & 1 & 0 & 0 & 0 & y & 0 & 0 & 1 & 1 \ 0 & 0 & 0 & 1 & 0 & 0 & 0 & y & 0 \end{bmatrix}$$

The weighted adjacency matrices, n = 3,  $x_0 + x_2$ ,  $(1 + y)I - W_{F,n}$ 



and the determinant from the Proposition:  $d_{3,F} = y + y^3$ We call these determinants Fryers' polynomials. They describe the process of generating all de Bruijn sequences (or all corresponding feedback functions) when starting from a non-singular feedback shift register.

## Spectral Theory of Graphs

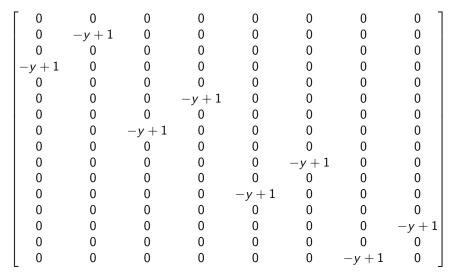
- Charles Delorme and Jean-Pierre Tillich
   The Spectrum of de Bruijn and Kautz Graphs
   Europ. J. Combinatorics, 19(1998), pp. 307-319.
- They have given a complete description of the spectrum of de Bruijn graph in terms of Tchebychev polynomials.
- We try to get similar results for weighted adjacency matrices related to Feedback Shift Registers.
- The first step is to diagonalize the weighted adjacency matrices.
- We have done a partial diagonalization.

n = 4

## $x_0 + x_1$ the matrix (1/16) H W H, H – Hadamard matrix The weighted adjacency matrix after *partial* diagonalization

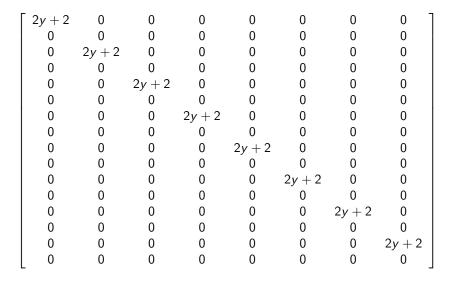
$\int y + 1$	0	0	0	0	0	0	0 ]
0	0	0	0	0	0	0	0
0	y + 1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	y + 1	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	y + 1	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	y + 1	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	y + 1	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	y + 1	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	y+1
0	0	0	0	0	0	0	0 ]

n = 4  $x_0 + x_1$  the matrix (1/16) H W H



 $y + 7y^3 + 7y^5 + y^7$ 

n = 4,  $x_0 + x_1 + x_2 + x_1 x_2$ , (1/8) H W H



 $x_0 + x_1 + x_2 + x_1 x_2$ 

# THANK YOU !