On The Nonlinearity of Maximum-length NFSR Feedbacks

Meltem Sönmez Turan

National Institute of Standards and Technology
meltem.turan@nist.gov

Abstract. Linear Feedback Shift Registers (LFSRs) are the main building block of many classical stream ciphers; however due to their inherent linearity, most of the LFSR-based designs do not offer the desired security levels. In the last decade, using Nonlinear Feedback Shift Registers (NFSRs) in stream ciphers became very popular. However, the theory of NFSRs is not well-understood, and there is no efficient method that constructs a cryptographically strong feedback function with maximum period and also, given a feedback function it is hard to predict the period. In this paper, we study the maximum-length NFSRs, focusing on the nonlinearity of their feedback functions. First, we provide some upper bounds on the nonlinearity of the maximum-length feedback functions, and then we study the feedback functions having nonlinearity 2 in detail. We also show some techniques to improve the nonlinearity of a given feedback function using cross-joining.

1 Introduction

Feedback Shift Registers (FSRs) are commonly used in stream cipher designs, due to their efficiency, large period and good statistical properties. FSRs with linear feedback function, Linear Feedback Shift Registers (LFSRs) are widely studied in the literature [1] and it is easy to find LFSRs with maximum period, $2^n - 1$, for a given length n. However, one important drawback of LFSR outputs is that they are completely linear, thus cryptographically insecure. Whenever $2n$ bits of the output is given, the sequence is totally predictable using the Berlekamp-Massey algorithm.

Various design attempts have been made to add nonlinearity to the ciphers based on LFSRs, such as combining outputs of several LFSRs using a nonlinear function, nonlinearly filtering the LFSR state or irregularly decimating the output [2]. However, most of these approaches do not offer desired security levels [3]. Due to the limitations of LFSRs, use of Nonlinear Feedback Shift Registers (NFSRs) became very popular. The eSTREAM Stream Cipher Project hardware finalists Grain [4], Mickey [5] and Trivium [6] use NFSRs as their main building blocks.

NFSRs constitute a larger class compared to LFSRs and they are more resistant to algebraic attacks. However the theory of NFSRs is not well-understood.

There is no efficient method that finds a cryptographically strong feedback function with maximum period 2^n for a given n and also, given a feedback function it is hard to predict the period.

Golomb presented a method to construct maximum-length NFSRs using primitive polynomials [1] (p. 115), however, these feedback functions have very low nonlinearity which allows them to be approximated using affine functions. Also, in 1982, Fredricksen [7] presented a survey on maximum-length NFSRs including construction methods and some properties. Tsueda et al. [8] proposed feedback-limited NFSRs and studied their properties in terms of correlation and linear complexity measures. Çalık et al. [9] studied maximum-length NFSRs focusing on the number of monomials.

In this paper, we study the nonlinearity of the feedback functions that generate maximum-length sequences. First, we provide some upper bounds on the nonlinearity of the feedback functions, and then we study the feedback functions having nonlinearity 2 in detail. We also show some techniques to increase the nonlinearity of a given feedback function using cross-joining and present some results on the relation between number of monomials and nonlinearity for extreme cases.

The paper is organized as follows. In Section 2, we give a brief introduction to Boolean functions and FSRs. The properties of maximum-length NFSRs are provided in Section 3. Section 4 focuses on the nonlinearity of feedback functions. Section 5 concludes the paper.

2 Preliminaries

2.1 Boolean Functions

A *Boolean function* f with n variables is a mapping from $\mathcal{F}_2^n$ to $\mathcal{F}_2$. Let α_i be the n-bit vector corresponding to the binary representation of integers $i = 0, 1, 2, \ldots, 2^n - 1$. For a Boolean function with n variables, the sequence

$$(f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1})) \tag{1}$$

is called the *truth table* of f. *Algebraic normal form* (ANF) of a Boolean function is the polynomial

$$f(x_1, x_2, \ldots, x_n) = c_0 \oplus c_1 x_1 \oplus \ldots \oplus c_{12\ldots n} x_1 x_2 \ldots x_n \tag{2}$$

with $c_{i_1 \ldots i_k}$'s in $\mathcal{F}_2$. The highest number of terms in a monomial with nonzero coefficient is called the *degree* of f. The Boolean functions with degree 1 are called *affine* and in particular for $c_0 = 0$, the functions are called *linear*.

The distance between two Boolean functions f and g is defined as the number of different entries in their truth table and denoted by $d(f, g)$, i.e., the weight of $f \oplus g$. Walsh transform of f is defined to be

$$W_f(\alpha) = \sum_{x \in F_2^n} (-1)^{f(x) \oplus \alpha . x}.$$

Nonlinearity of a Boolean function f, denoted as $Nl(f)$ is the minimum distance of f to the set of all affine functions, which is

$$2^{n-1} - \frac{1}{2} \max_{\alpha} \{|W_f(\alpha)|\}. \tag{3}$$

Nonlinearity of a Boolean function is bounded by $2^{n-1} - 2^{n/2-1}$. The Boolean functions with even number of variables that achieve this bound are called *bent functions*. Weight of bent functions can take two values $2^{n-1} \pm 2^{n/2-1}$, i.e., they are not balanced, thus not very useful in cryptographic applications.

Two n-bit Boolean functions $f(x)$ and $h(x)$ are called *affine equivalent*, if $h(x) = f(Ax + b)$, where A is an n by n binary non-singular matrix and b is an n-bit binary vector and it is known that $Nl(f)$ and $Nl(h)$ are equal.

2.2 Feedback Shift Registers

A *FSR* is a device that shifts its contents into adjacent positions within the register and fills the position on the other end with a new value generated by the *feedback function*. The individual delay cells of the register are called the *stages* and the number of the stages n is called the *length* of FSR. The contents of the n stages are called the *state* of the FSR. The n bit vector $(s_0, s_1, \ldots, s_{n-1})$ initially loaded into FSR state specify the *initial state*. A block diagram of a FSR is given in Figure 1.

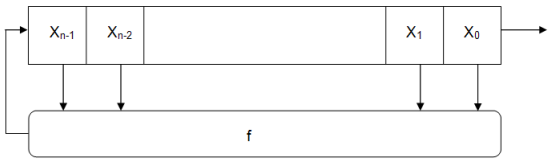

Fig. 1. Block diagram of a Feedback Shift Register.

A FSR is uniquely determined by its length n and n variable Boolean feedback function $f(x_1, x_2, \ldots, x_n)$. The output sequence $\mathbf{S}= \{s_0, s_1, s_2, \ldots\}$ of a FSR satisfy the following recursion

$$s_{n+i} = f(s_i, \ldots, s_{n-1+i}),\ i \geq 0 \tag{4}$$

given the initial state $(s_0, s_1, \ldots, s_{n-1})$.

For LFSRs, this recursion is linear and may be represented using the *characteristic polynomial*,

$$C(x) = \sum_{i=0}^{n} c_i x^{n-i} \tag{5}$$

with $c_0 = 1$. If $C(x)$ is a primitive polynomial with degree n, then each of the $2^n - 1$ non-zero initial states of the LFSR produce an output with maximum possible period $2^n - 1$. Outputs of maximum-length LFSRs are called *maximal length sequences* or *m-sequences*.

Let $\mathcal{L}_n$ be the set of all linear feedback functions that generate m-sequence. The number of primitive polynomial of degree n over $\mathcal{F}_2$ is given by $\phi(2^n-1)/n$, where $\phi(n)$ is the Euler's phi function, hence

$$|\mathcal{L}_n| = \phi(2^n - 1)/n. \tag{6}$$

For NFSRs, the output sequences can achieve the period of 2^n. Such sequences include each n-bit pattern exactly once and are called *de Bruijn sequences*. Let $\mathcal{D}_n$ be the set of all feedback functions that generate de Bruijn sequence and

$$|\mathcal{D}_n| = 2^{2^{n-1}-n} \text{ [10]}. \tag{7}$$

3 NFSRs and de Bruijn Sequences

3.1 Properties of Maximum-Length NFSRs

In this part of the study, we survey some of the necessary conditions of the feedback function $f(x_1, \ldots, x_n)$ to generate de Bruijn sequences.

To guarantee that every state has a unique predecessor and successor, f should be written in the form $f(x_1, \ldots, x_n) = x_1 + g(x_2, \ldots, x_n)$ [1]. Some necessary conditions on f and g to generate a de Bruijn sequence are given as follows;

1. To avoid all zero cycle, $f(0, \ldots, 0) = 1$, i.e. $c_0 = 1$. Due to unique predecessor and successor property, $f(1, 0, \ldots, 0) = 0$.
2. To avoid all one cycle, $f(1, \ldots, 1) = 0$, therefore the number of monomials in f is even. Due to unique predecessor and successor property, $f(0, 1, \ldots, 1) = 1$.
3. To avoid the cycle $(00 \ldots 01)$, there exists a coefficient $c_i = 0$, for $i = 2, \ldots, n$ [11].
4. The parity of the cycles generated by a FSR is equal to the parity of the truth table of g [1]. To achieve one maximum-length cycle, parity of the truth table of g should be 1, which implies $c_{23\ldots n} = 1$.
5. The weight $w(g)$ of g satisfies the following inequality

$$Z_{n-1} \leq w(g) \leq 2^{n-1} - Z_n^* + 1 \tag{8}$$

   where Z_n is $\frac{1}{n}\sum_{d|n} \phi(d)2^{n/d}$ and Z_n^* is $\frac{Z_n}{2} - \frac{1}{2n}\sum \phi(2d)2^{n/2d}$ with summation over all even divisors of n [7].
6. Let $f = x_1 + g(x_2, \ldots, x_n)$ generate a de Bruijn sequence and $n > 2$. Then,

$$g(x_2, \ldots, x_n) \neq g(x_n, \ldots, x_2), \tag{9}$$

   i.e., g is not rotation symmetric [9].

4 On The Nonlinearity of Feedback Functions

In this part of the paper, we study on the nonlinearity of the feedback functions that generate de Bruijn sequences.

Proposition 1. *Let $f(x_1, \ldots, x_n) \in D_n$. The nonlinearity of f satisfies*

$$Nl(f) \equiv 2 \ (mod\ 4).$$

Proof. Due to the unique predecessor and successor property, f has the form; $f = x_1 + g$. For any linear function l, $g \oplus l$ includes the monomial $x_2 \cdots x_n$, therefore weight of $g \oplus l$ is always odd. This guarantees that nonlinearity of g is odd. Let $Nl(g) = 2k + 1$, then $Nl(f) = 2(2k + 1) = 4k + 2$, hence $Nl(f) \equiv 2 \ (mod\ 4)$.

Next, we show some upper bounds on the nonlinearity of $f \in \mathcal{D}_n$. As mentioned in Section 2, nonlinearity of an n-variable Boolean function is bounded by $2^{n-1} - 2^{n/2-1}$. Since the functions achieving this bound are not balanced, they cannot generate de Bruijn sequences. Next proposition provides an upper bound on the nonlinearity of $f \in \mathcal{D}_n$ based on the weight of g, $w(g)$.

Proposition 2. *The nonlinearity of $f \in \mathcal{D}_n$ is bounded by*

$$Nl(f) \ \leq \min\{2^n - 2Z_n^* + 2, 2^n - 2Z_{n-1}\}.$$

Proof. According to [7], the weight of g satisfies the following inequality

$$Z_{n-1} \leq w(g) \leq 2^{n-1} - Z_n^* + 1.$$

Since the distance between g and the constant zero function is $w(g)$, $Nl(g) \leq w(g)$ and $Nl(f) \ \leq \ 2w(g)$. Then

$$Nl(f) \leq 2^n - 2Z_n^* + 2 \tag{10}$$

is satisfied. Similarly, the distance between g and the constant one function is equal to $2^{n-1} - w(g)$. Then,

$$Nl(f) \leq 2^n - 2Z_{n-1} \tag{11}$$

is satisfied. Combining (10) and (11), nonlinearity of f is bounded by

$$Nl(f) \ \leq \min\{2^n - 2Z_n^* + 2, 2^n - 2Z_{n-1}\}.$$

It should be noted these bounds are not tight, as they are only based on the weight of g, not on the location of 1's in the truth table of g. Another bound for the nonlinearity of f is provided in the next proposition.

Proposition 3. *The nonlinearity of $f \in \mathcal{D}_n$ is bounded by*

$$2^{n-1} - 2^{(n-1)/2},$$

for $n > 2$.

Proof. The nonlinearity of the (n-1)-variable Boolean function g is bounded by $2^{(n-2)}-2^{(n-3)/2}$*. It is known that for* $n > 2$*, degree of an n-variable bent function is bounded by* $n/2$*. Since the degree of g is* $n-1$*, this bound cannot be achieved by g. Therefore, the following is satisfied;*

$$\begin{aligned} Nl(f) &< 2(2^{(n-2)} - 2^{(n-3)/2}) \\ &< 2^{n-1} - 2^{(n-1)/2}. \end{aligned}$$

Table 1 shows two upper bounds on the nonlinearity of f. The first bound is the generic nonlinearity bound for Boolean functions, whereas the second bound is obtained using Proposition 1 and 3.

n	Bound 1	Bound 2
5	13	10
6	28	26
7	58	54
8	120	114
9	244	238
10	496	486

Table 1. Nonlinearity bounds on the $Nl(f)$.

Let $\gamma_n(c)$ denote the number of n-variable maximum-length feedback functions with nonlinearity c. Trivially, $\sum_c \gamma_n(c) = 2^{2^{n-1}-n}$. The value of $\gamma_n(c)$ for arbitrary n and c is not known. According to the Proposition 1, $\gamma_n(4k) = \gamma_n(4k+1) = \gamma_n(4k+3) = 0$, $k \geq 0$. Table 2 shows the distribution of $\gamma_n(c)$ for $n = 5, 6$.

Proposition 4. $\gamma_n(c) \equiv 0$ *(mod 4), for even* $n \geq 3$*.*

Proof. The complement $\mathbf{c}S$ *and the reverse* $\mathbf{r}S$ *of a de Bruijn sequence* $S = \{s_0, s_1, \ldots, s_{2^n-1}\}$ *is defined to be* $cS = \{s'_0, s'_1, \ldots, s'_{2^n-1}\}$ *where* s'_i *is the binary complement of* s_i *and* $rS = \{s_{2^n-1}, \ldots, s_1, s_0\}$*. Let* $f(x_1, x_2, \ldots, x_n) \in \mathcal{D}_n$ *generate S, then* $f(x_1, x'_2, \ldots, x'_n)$ *generates* $\mathbf{c}S$ *and* $f(x_1, x_n, x_{n-1}, \ldots, x_2)$ *generates* $\mathbf{r}S$ *[7]. Since* $f(x_1, \ldots, x_n)$*,* $f(x_1, x'_2, \ldots, x'_n)$*,* $f(x_1, x_n, \ldots, x_2)$ *and* $f(x_1, x'_n, x'_{n-1}, \ldots, x'_2)$ *are affine equivalent, the nonlinearity of feedback functions generating* $S, \mathbf{r}S, \mathbf{c}S$ *and* $\mathbf{rs}S$ *are equal. In [12], it is shown that* $S, \mathbf{r}S, \mathbf{c}S$ *and* $\mathbf{rs}S$ *are distinct sequences, for even n. Hence,* $\gamma_n(c) \equiv 0$ *(mod 4) for even* $n \geq 3$*.*

4.1 Feedback Functions with Nonlinearity 2

As shown in Proposition 1, the minimum value of the nonlinearity of a maximum-length feedback function is 2. Following proposition shows a construction method of feedback functions having nonlinearity 2 using primitive polynomials.

n	c	$\gamma_n(c)$
5	2	24
	6	1128
	10	896
6	2	32
	6	7408
	10	352,752
	14	6,491,072
	18	42,601,512
	22	17,656,088

Table 2. The distribution of $\gamma_n(c)$ for $n = 5, 6$.

Proposition 5. *Let $F = c_1x^n + c_2x^{n-1} + \ldots + c_nx + 1$ be a primitive polynomial with degree n over the finite field GF(2).*

(i) $f_1(x_1, \ldots, x_n) = c_1x_1 + c_2x_2 + \ldots + c_nx_n + x_2'x_3' \cdots x_n' \in \mathcal{D}_n$ *[1] and* $Nl(f_1) = 2$.

(ii) $f_2(x_1, \ldots, x_n) = 1 + c_1x_1 + c_2x_2 + \ldots + c_nx_n + x_2x_3 \cdots x_n \in \mathcal{D}_n$ *[9] and* $Nl(f_2) = 2$.

Proof. Given a primitive polynomial $F(x)$, it is known that $l(x) = c_1x_1 + \ldots + c_nx_n$ generates a state diagram with two cycles, one of which is the all-zero cycle and the other is a cycle of length $2^n - 1$.

(i) These two cycles can be combined by adding $x_2'x_3' \cdots x_n'$ to $l(x)$, which is equivalent to changing the the truth table entries corresponding to $(10 \ldots 0)$ and $(00 \ldots 0)$ [1]. Then, $f_1(x_1, \ldots, x_n) = l(x) + x_2'x_3' \cdots x_n'$ generates a cycle with length 2^n and since only two truth table entries of l are changed, $Nl(f_1) = 2$.

(ii) Since $F(x)$ is primitive, the number of monomials in $F(x)$, also in $l(x)$, is odd, so $l'(x) = c_1x_1' + \ldots + c_nx_n' = 1 + c_1x_1 + c_2x_2 + \ldots + c_nx_n$. $l'(x)$ generates two cycles one of which is the all-one cycle and the other one is the cycle of length $2^n - 1$, which are the complements of the cycles generated by $l(x)$. These two cycles are combined by adding $x_2x_3 \cdots x_n$ to $f(x)$, which is equivalent to changing the truth table entries corresponding to $(011 \ldots 1)$ and $(11 \ldots 1)$, therefore $Nl(f_2) = 2$.

Table 3 lists 12 feedback functions generated using the primitive polynomials, for $n = 5$. There exists other feedback functions with nonlinearity 2 which can be constructed using affine functions having two cycles, where length of the shortest cycle is greater than 1. These functions are highlighted in Table 4.1. These cycles can also be combined in various ways to generate feedback functions having nonlinearity 2, as given in the next example.

Primitive polynomial	Feedback Functions with Nonlinearity 2
$1+x^2+x^5$	$x_1+x_4+x_2'x_3'x_4'x_5'$ $1+x_1+x_4+x_2x_3x_4x_5$
$1+x+x^2+x^3+x^5$	$x_1+x_3+x_4+x_5+x_2'x_3'x_4'x_5'$ $1+x_1+x_3+x_4+x_5+x_2x_3x_4x_5$
$1+x^3+x^5$	$x_1+x_3+x_2'x_3'x_4'x_5'$ $1+x_1+x_3+x_2x_3x_4x_5$
$1+x+x^3+x^4+x^5$	$x_1+x_2+x_3+x_5+x_2'x_3'x_4'x_5'$ $1+x_1+x_2+x_3+x_5+x_2x_3x_4x_5$
$1+x^2+x^3+x^4+x^5$	$x_1+x_2+x_3+x_4+x_2'x_3'x_4'x_5'$ $1+x_1+x_2+x_3+x_4+x_2x_3x_4x_5$
$1+x+x^2+x^4+x^5$	$x_1+x_2+x_4+x_5+x_2'x_3'x_4'x_5'$ $1+x_1+x_2+x_4+x_5+x_2x_3x_4x_5$

Table 3. Construction of f with nonlinearity 2, for n=5, using primitive polynomials.

Example 1. The cycle decomposition of affine function $f = 1 + x_1 + x_3 + x_5$ includes two cycles; (00000101101110011111010010011) and (01). The long cycle includes all n-bit patterns except (10101) and (01010). The (01) pattern is embedded to the long cycle such that the new generated cycle includes all n-bit patterns and it can be done in two ways as shown in the following figure.

New sequence	Feedback function
(00000101**01**10111001111101001000 11)	$f + (x_3 + x_5)\, x_2'\, x_4'$
(00001011011100111111010**01**001000110)	$f + (x_2 + x_4)\, x_3'\, x_5'$

Table 4. Combining the two cycles of $f = 1 + x_1 + x_3 + x_5$

Proposition 6. *The number of maximum-length feedback functions with nonlinearity 2 satisfies*

$$2\frac{\phi(2^n-1)}{n} \leq \gamma_n(2) \leq 2^{2n}, \tag{12}$$

where $\phi(n)$ is the Euler's phi function.

Proof. Proposition 5 provides a method to construct two distinct feedback functions with nonlinearity 2 using a primitive polynomial of degree n. Therefore,

Affine function	Cycle decomposition	# cycles
$1+x_1+x_2+x_3+x_5$	(00000101110110101001111000110 01) (1)	2
$1+x_1+x_2+x_3+x_4$	(00000110110011110100101011100 01) (1)	2
$1+x_1+x_4$	(00000110010110111101010001001 11) (1)	2
$1+x_1+x_3$	(00000111001000101011110110100 11) (1)	2
$1+x_1+x_3+x_4+x_5$	(00000100011101010010111100110 11) (1)	2
$1+x_1+x_2+x_4+x_5$	(00000100110001111001010110111 01) (1)	2
$1+x_1+x_3+x_5$	(00000101101110011111010010001 1) (01)	2
$1+x_1+x_2+x_4$	(00000110001001011111001110110 1) (01)	2
$1+x_1+x_2+x_3$	(00000111010110111111000101001) (0011)	2
$1+x_1+x_4+x_5$	(00000100101000111111011010111) (0011)	2
$1+x_1$	(0000011111) (0001011101) (0010011011) (01)	4
$1+x_1+x_5$	(000001010110011101111) (0001101) (001) (1)	4
$1+x_1+x_3+x_4$	(000001101011) (0001) (001010011111) (0111)	4
$1+x_1+x_2$	(000001111011100110101) (0001011) (001) (1)	4
$1+x_1+x_2+x_5$	(00000101) (00011011) (00100111) (01011111)	4
$1+x_1+x_2+x_3+x_4+x_5$	(000001) (000111) (001011) (001101) (01) (011111)	6

Table 5. Cycle decompositions of affine functions with $c_0 = 1$, for n=5.

$\gamma_n(2)$ is lower bounded by 2 times the number of primitive polynomials, which is equal to $2\frac{\phi(2^n-1)}{n}$.

The number of Boolean functions having nonlinearity 2 is calculated by counting the number of 2 possible changes to all affine functions and is equal to $\binom{2^n}{2}2^{n+1}$. For the functions in $\mathcal{D}_n$, the changes should be symmetric, in other words, $f(a_1, a_2, \ldots, a_n)$ and $f(a_1+1, a_2, \ldots, a_n)$ should be changed simultaneously. Therefore, the number of feedback functions with nonlinearity 2 is upper bounded by $\binom{2^{n-1}}{1}2^{n+1} = 2^{2n}$.

Proposition 7. *The number of maximum-length feedback functions with nonlinearity t is bounded by*

$$\gamma_n(t) \leq \binom{2^{n-1}}{t/2} 2^{n+1}, \tag{13}$$

for even $t < 2^{n-2}$.

Proof. The number of functions having nonlinearity t is calculated by counting the number of t possible changes to all affine functions and is equal to $\binom{2^n}{t}2^{n+1}$, for $t < 2^{n-2}$ [13]. Since the changes should be symmetric, the total number of changes that can be given to an affine function is bounded by $\binom{2^{n-1}-2}{t/2}$.

4.2 Cross-joining to Increase Nonlinearity

Cross-joining is a well-known method to construct maximum-length feedback functions given another feedback function [7]. Given $f \in \mathcal{D}_n$, the method first flips one position of the truth table of g, this splits the output of f into two cycles. Then, if a second position in the truth table of g exists, such that flipping it combines the two cycles to produce a new cycle of length 2^n, then this pair of positions is called a *cross-join pair*.

In the previous section, construction of feedback functions with nonlinearity 2 is described. By cross-joining, the nonlinearity of the feedback functions can be improved. Cross-joining flips four positions of the truth tables, therefore the nonlinearity of the newly constructed feedback function is bounded by $Nl(f)+4$. Helleseth and Kløve [14] proved that the number of cross-join pairs in an n-bit maximum-length LFSR is $(2^{n-1}-1)(2^{n-1}-2)/6$.

Cross-joining is also equivalent to dividing a de Bruijn sequence into five parts and permuting the parts as given in Figure 2. The permutation basically interchanges the positions of the part II and part IV. It is possible to verify that the sequence generated by interchanging the positions of the unique runs of n and $n-2$ zeros in a de Bruijn sequence, is also a de Bruijn sequence. This is also true for interchanging the unique runs of n and $n-2$ ones [12].

$$(\leftarrow I \rightarrow || \leftarrow II \rightarrow || \leftarrow III \rightarrow || \leftarrow IV \rightarrow || \leftarrow V \rightarrow)$$
$$\Longleftrightarrow$$
$$(\leftarrow I \rightarrow || \leftarrow IV \rightarrow || \leftarrow III \rightarrow || \leftarrow II \rightarrow || \leftarrow V \rightarrow)$$

Fig. 2. Cross join pair effect on de Bruijn sequence

Proposition 8. *If $f(x_1, \ldots, x_n) \in \mathcal{D}_n$, then $f_1 = f \oplus (x_2 \oplus x_n)x_3'x_4' \ldots x_{n-1}'$ and $f_2 = f \oplus (x_2 \oplus x_n)x_3x_4 \ldots x_{n-1} \in \mathcal{D}_n$.*

Proof. The monomial $(x_2 \oplus x_n)x_3'x_4' \ldots x_{n-1}'$ takes the value 1 in four positions of the truth table of f, i.e. flips the output of f in the following inputs; $f(0,1,0,\ldots,0)$, $f(1,1,0,\ldots,0)$, $f(0,\ldots,0,1)$, $f(1,0,\ldots,0,1)$. These changes results in interchanging the runs of n and $n-2$ zeros. Similarly, adding $(x_2 \oplus x_n)x_3x_4 \ldots x_{n-1}$, changes the following outputs of f; $f(0,1,\ldots,1,0)$, $f(1,1,\ldots,1,0)$, $f(0,0,1\ldots,1)$, $f(1,0,1,\ldots,1)$. These changes results in interchanging the runs of n and $n-2$ ones. Therefore, f_1 and f_2 also generate de Bruijn sequences.

Given a feedback function with nonlinearity 2, applying the changes given in Proposition 8 increases the nonlinearity to 6. Independent cross-join pairs may be applied to further increase the nonlinearity of the feedback function.

4.3 Number of Monomials

Hardware efficiency of NFSRs is extremely important, especially for stream ciphers designed for restricted environments. In general, the feedback functions with less number of monomials are implemented more efficiently in hardware and these functions are of interest. Çalık et al. [9] studied the number of monomials in the maximum-length feedback functions and showed that it is at least 4.

Proposition 9. *Let $f \in \mathcal{D}_n$ be a 4-monomial feedback function.*

$$Nl(f) \leq \frac{2^n}{n-1}.$$

Proof. 4-monomial feedback functions are of the form;

$$f(x_1, x_2, \ldots, x_n) = 1 + x_1 + x_2 \cdots x_n + h(x_2, \ldots, x_n),$$

where the degree of $h(x)$ is upper bounded by $log_2 n$ [9]. Nonlinearity of f is equal to $2Nl(h(x) + x_2 \cdots x_n)$ and the weight of $h(x) + x_2 \cdots x_n$, which is equal to $2^{n-1-deg(h(x))} - 1$, gives two upper bounds on the nonlinearity.

$$\begin{aligned} Nl(h(x) + x_2 \cdots x_n) &\leq \min\{2^{n-1-deg(h(x))} - 1, 2^{n-1}(1 - 2^{-deg(h(x)}) + 1\} \\ &\leq \min\{2^{n-1-log_2 n} - 1, 2^{n-1}(1 - 2^{log_2 n}) + 1\} \\ &\leq \min\{2^{n-1}/n - 1, 2^{n-1}(1 - 1/n) + 1\} \\ &\leq 2^{n-1}/n - 1. \end{aligned}$$

Then, $Nl(f) \leq \frac{2^n}{n-1}$ is satisfied.

Çalık et al. [9] conjectured that for $n > 12$, the degree of $h(x)$ is 1, after experimenting for $n \leq 36$. Assuming the conjecture is true, we can say that the nonlinearity of 4-monomial feedback functions is 2 for $n > 12$.

The maximum number of monomials in a maximum-length feedback function is $2^n - 1$ and these functions are of the form

$$f = x_1 + x_2' \cdots x_n' + x_i,$$

for $2 \leq i \leq n$ [9]. It is easy to verify that the nonlinearity of these functions is also 2.

5 Conclusion

To have a better understanding of NFSRs, in this study, we focused on the nonlinearity properties of feedback functions of maximum-length NFSRs. We provided some upper bounds on the nonlinearity, and studied the feedback functions having nonlinearity 2 and provided some techniques to increase the nonlinearity of a given feedback function using cross-joining.

References

1. S. W. Golomb. *Shift Register Sequences*. Holden-Day, Inc., Laguna Hills, CA, USA, 1967.
2. A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, Boca Raton, Florida, 1996.
3. A. Braeken and J. Lano. On the (Im)Possibility of Practical and Secure Nonlinear Filters and Combiners. In *Selected Areas in Cryptography*, pages 159–174, 2005.
4. M. Hell, T. Johansson, and W. Meier. Grain - A Stream Cipher for Constrained Environments. eSTREAM, ECRYPT Stream Cipher Project, Report 2005/010, 2005.
5. S. Babbage and M. Dodd. The Stream Cipher MICKEY (version 1). eSTREAM, ECRYPT Stream Cipher Project, Report 2005/015, 2005.
6. C. De Cannière and B. Preneel. Trivium Specifications. eSTREAM, ECRYPT Stream Cipher Project, Report 2005/030, 2005.
7. H. Fredricksen. A Survey of Full Length Nonlinear Shift Register Cycle Algorithms. 24(2):195–221, 1982.
8. A. Tsuneda, K. Kudo, D. Yoshioka, and T. Inoue. Maximal-Period Sequences Generated by Feedback-Limited Nonlinear Shift Registers. *IEICE Transactions*, 90-A(10):2079–2084, 2007.
9. Ç. Çalık, M. Sönmez Turan, and F. Özbudak. On Feedback Functions of Maximum Length Nonlinear Feedback Shift Registers. *IEICE Transactions*, 93-A(6):1226–1231, 2010.
10. N. G. de Bruijn. A Combinatorial Problem. *In Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen. Series A*, 49(7):758–764, 1946.
11. R. Gonzalo, D. Ferrero, and M. Soriano. Some Properties of Non Linear Feedback Shift Registers with Maximum Period. *Proc. Sixth Int. Conf. Telecommunications Systems*, 1998.
12. T. Etzion and A. Lempel. On the Distribution of de Bruijn Sequences of Given Complexity. *IEEE Transactions on Information Theory*, 30(4):611–614, 1984.
13. C.K. Wu. Distribution of Boolean Functions with Nonlinearity $2^{(n-2)}$. In *Proceedings of ChinaCrypt'94*, pages 10–14, China, 1994. Springer-Verlag.
14. T. Helleseth and T. Kløve. The Number of Cross-join Pairs in Maximum Length Linear Sequences. *IEEE Transactions on Information Theory*, 37(6):1731–1733, 1991.