# On cross joining de Bruijn sequences

Johannes Mykkeltveit and Janusz Szmidt

Abstract. We explain the origins of Boolean feedback functions of nonlinear feedback shift registers (*NLFSRs*) of fixed order $n$ generating de Bruijn binary sequences. They all come into existence by cross joining operations starting from one maximum period feedback shift register, e.g., a linear one which always exists for any order $n$. The result obtained yields some constructions of *NLFSRs* generating maximum period $2^n - 1$ binary sequences.

## 1. Introduction

The task of this note is to get insight into construction of Boolean feedback functions of *NLFSRs* generating maximum period binary sequences. At the International Workshop on Coding and Cryptography 2013 in Bergen we discussed the problem whether it was true or not that an arbitrary de Bruijn sequence could be obtained by applying cross-join pair operations to a given one. It seems that this has been a several decades old unsolved problem. In this note we solve this problem in the affirmative. In terms of Boolean feedback functions this result indicates which algebraic operations must be applied to get all feedback functions generating de Bruijn sequences starting from a given one. The same results are true when one deals with modified de Bruijn sequences of period $2^n - 1$ and the corresponding feedback functions.

Feedback shift registers (*FSRs*) are useful in generating periodic sequences, and that is the task for which they are mostly used in communication and cryptographic systems. Linear feedback shift registers (*LFSRs*) and *NLFSRs* are the main buiding blocks of many stream ciphers. The *LFSRs* are well-understood mathematically. The investigation of *NLFSRs* started in the pioneering book of Golomb [**9**] and has continued for several decades. In cryptographic applications, *NLFSRs* generating modified de Bruijn sequences are important since in special cases the algebraic normal form (*ANF*) of the corresponding Boolean feedback functions is simpler than that of de Bruijn sequences (see, e.g., [**16, 8**]).

The operation of joining and disjoining cycles generated by nonsingular *FSRs* was discussed in Golomb's book [**9**]. After that the notion of cross-join pairs was employed to construct new *NLFSRs* from given ones (see, e.g. [**14, 7, 6, 17, 12, 20**]). Helleseth and Kløve [**11**] proved an important result which gives the

1991 *Mathematics Subject Classification.* Primary 94A55, 94C10; Secondary 06E30, 12Y05.

*Key words and phrases.* De Bruijn sequences, cross-join pairs, nonlinear feedback shift registers, maximum period, quadratic $m$-sequences.

number of cross-join pairs for a binary $m$-sequence. In a recent paper of Dubrova [**4**] cross-join pairs were used to construct Galois *NLFSRs* with maximum period.

Methods for finding *NLFSRs* with simple *ANF* of the feedback function were presented in [**8, 2, 19, 3**]. Gong and Mandal [**15**] following Mykkeltveit *et al.* [**18**] developed a recursive method for constructing maximum period *NLFSRs*. Chan, Games and Rushanan [**2**] conjectured existence of quadratic $m$-sequences for each order $n$. In [**3**] we have verified this conjecture experimentally up to order $n = 29$, finding simple quadratic *NLFSRs* by searching methods. In the present note we aim to relate the existence of *NLFSRs* generating quadratic $m$-sequences to the construction of cross joining. We give an example for order $n = 7$ and formulate necessary conditions to be able to construct maximum period *NLFSRs* whose feedback functions have some properties.

The paper is organized as follows. Section 2 recalls the known definitions and theorems about de Bruijn sequences and nonlinear feedback shift registers. Section 3 contains our main theorem together with its proof. Quadratic $m$-sequences and examples of *NLFSRs* of order 4 and 7 generating modified de Bruijn sequences are presented in Section 4. In that section we relate the Chan, Games and Rushanan conjecture [**2**] to existence of a suitable collection of cross-join pairs for a given $m$-sequence.

## 2. Some definitions and known theorems

Let $\mathbb{F}_2 = \{0, 1\}$ denote the binary field and $\mathbb{F}_2^n$ the vector space of all binary $n$-tuples. A binary feedback shift register (*FSR*) of order $n$ is a mapping

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

of the form

$$\mathfrak{F} : (x_0, x_1, \ldots, x_{n-1}) \longmapsto (x_1, x_2, \ldots, x_{n-1}, f(x_0, x_1, \ldots, x_{n-1})), \tag{2.1}$$

where the *feedback function* $f$ is a Boolean function of $n$ variables. The *FSR* is called *non-singular* if the mapping $\mathfrak{F}$ is one-to-one, i.e., $\mathfrak{F}$ is a bijection of $\mathbb{F}_2^n$. The *FSR* is called linear (*LFSR*) if the feedback function $f$ is linear, and nonlinear (*NLFSR*) if $f$ is nonlinear, i.e., $f$ has higher degree terms in its algebraic normal form (*ANF*).

Definition 2.1. A *de Bruijn sequence* of order $n$ is a sequence of length $2^n$ of elements of $\mathbb{F}_2$ in which all different $n$-tuples appear exactly once.

It was proved by Flye Sainte-Marie [**5**] in 1894 and independently by de Bruijn [**1**] in 1946 that the number of cyclically non-equivalent sequences satisfying Definition 2.1 is equal to

$$B_n = 2^{2^{n-1}-n}. \tag{2.2}$$

Definition 2.2. A *modified* de Bruijn sequence of order $n$ is a sequence of length $2^n - 1$ obtained from a de Bruijn sequence of order $n$ by removing one zero from the tuple of $n$ consecutive zeros.

Let $(s_t) = (s_0, s_1, \cdots, s_{2^n-1})$ be a de Bruijn sequence. For the purpose of this note we put $S_i = (s_i, s_{i+1}, \cdots, s_{i+(n-2)})$, and write the de Bruijn sequence as $(S_t) = (S_0, S_1, \cdots, S_{2^n-1})$ (indexes are reduced mod $2^n$ in such a way that

$0 \le t \le 2^n - 1$). In the latter representation each $(n-1)$-vector occurs exactly twice.

Two elements $U, V \in \mathbb{F}_2^{n-1}$ constitute a *cross-join pair* for the sequence $(s_t)$ if one can shift $(S_t)$ cyclically so that $U, V$ occur in the order $U, \cdots, V, \cdots, U, \cdots, V$.

It follows that for the pairs of states of a FSR generating the sequence $(s_t)$: $\alpha = (u, U)$, $\widehat{\alpha} = (\overline{u}, U)$ and $\beta = (v, V)$, $\widehat{\beta} = (\overline{v}, V)$, where $\overline{u} = u + 1$ is the negation of the bit $u$, they occur in the order $\alpha$, $\beta$, $\widehat{\alpha}$, $\widehat{\beta}$.

THEOREM 2.3. *Let $(s_t)$ be a de Bruijn sequence of order $n$. Then there exists a Boolean function $F(x_1, \cdots, x_{n-1})$ such that*

$$s_{t+n} = s_t + F(s_{t+1}, \cdots, s_{t+n-1}), \quad t = 0, 1 \cdots, 2^n - n - 1. \tag{2.3}$$

PROOF. See [**9**] where a more general result is given. □

This means that for the feedback function of (2.1) we have

$$f(x_0, x_1, \cdots, x_{n-1}) = x_0 + F(x_1, \cdots, x_{n-1}) \tag{2.4}$$

and de Bruijn sequences are generated by some non-singular *FSRs*. The problem mentioned in the Introduction can be formulated as follows: describe all feedback functions of *FSRs* generating all de Bruijn sequences. As concerns applications in cryptography we are interested in modified de Bruijn sequences since some feedback functions of *NLFSRs* generating these sequences have simple *ANF*.

THEOREM 2.4. *Let $(s_t)$ be a de Bruijn sequence satisfying (2.3). Let $U, V$ be a cross-join pair for $(s_t)$ and let $G(x_1, \cdots, x_{n-1})$ be obtained from $F(x_1, \cdots, x_{n-1})$ by the negation of $F(U)$ and $F(V)$. Then $G(x_1, \cdots, x_{n-1})$ also generates a de Bruijn sequence $(u_t)$, say. We say that $(u_t)$ is obtained from $(s_t)$ by the cross-join pair operation.*

PROOF. The negation of $F(U)$ will split the de Bruijn sequence into two sequences, and the negation of $F(V)$ will join these two sequences again (since $U, V$ is a cross-join pair). □

## 3. The main theorem

THEOREM 3.1. *Let $(u_t), (v_t)$ be two de Bruijn sequences of order $n$. Then $(v_t)$ can be obtained from $(u_t)$ by repeated application of the cross-join pair operation.*

PROOF. First, we observe that the cross-join pair operation leads to an equivalence relation in the set of all de Bruijn sequences. We order the truth tables of the functions $F$ in (2.4) lexicographically and denote this ordered set by $S$. We choose the ordering in such a way that $F(0, 0, \cdots, 0)$ is the most significant digit. Let $T_1$ be the equivalence class containing the lexicographically largest de Bruijn sequence. Suppose that the theorem is false. Then there must exist a nonempty equivalence class $T_2$ different from $T_1$ and let $H$ be the truth table for the lexicographically largest de Bruijn sequence in $T_2$. Then $H$ has the following two properties:

1. It is not the lexicographically largest de Bruijn sequence.

2. Any cross-join pair operation which can be applied to $H$ will result in a truth table less than $H$.

Define:

$$S_1 = \{F \in S \ : \ F \le H\}, \quad S_2 = \{F \in S \ : \ F > H\}.$$

We are done if we can prove that $H$ does not exist. Let $K \in S_2$. Let $(z_1, \cdots, z_{n-1})$ be the smallest $(n-1)$-vector such that $H(z_1, \cdots, z_{n-1})$ is different from $K(z_1, \cdots, z_{n-1})$. Since $H < K$ we have $H(z_1, \cdots, z_{n-1}) = 0$ and $K(z_1, \cdots, z_{n-1}) = 1$ and the choice of $(z_1, \cdots, z_{n-1})$ implies that if

$$(u_1, \cdots, u_{n-1}) < (z_1, \cdots, z_{n-1})$$

then

$$H(u_1, \cdots, u_{n-1}) = K(u_1, \cdots, u_{n-1}).$$

Let $H1$ be obtained from $H$ by putting $H1(z_1, \cdots, z_{n-1}) = 1$ and keeping $H1 = H$ for all other function arguments. Clearly this change will split the de Bruijn sequence so that $H1$ generates two sequences $C_1$ and $C_2$, say.

We have

$$H1(z_1, \cdots, z_{n-1}) = K(z_1, \cdots, z_{n-1}), \tag{3.1}$$

which implies that $(z_0, z_1, \cdots, z_{n-1})$ and $(z_1, \cdots, z_{n-1}, z_0 + K(z_1, \cdots, z_{n-1}))$ either both belong to $C_1$ or both belong to $C_2$. It is no restriction to assume that they both belong to $C_1$.

Since $K$ generates a de Bruijn sequence there exists an $n$-tuple $(v_0, \cdots, v_{n-1})$ such that

$$(v_0, v_1, \cdots, v_{n-1}) \in C_1$$

and

$$(v_1, \cdots, v_{n-1}, v_0 + K(v_1, \cdots, v_{n-1}) \in C_2,$$

and since $H1$ generates $C1$ we have

$$(v_1, \cdots, v_{n-1}, v_0 + H1(v_1, \cdots, v_{n-1})) \in C_1.$$

Since we have (3.1) and $H1(v_1, \cdots, v_{n-1}) \neq K(v_1, \cdots, v_{n-1})$ this implies that

$$(z_1, \cdots, z_{n-1}) \neq (v_1, \cdots, v_{n-1}),$$

and since $(z_1, \cdots, z_{n-1})$ is the smallest $(n-1)$-tuple $(x_1, \cdots, x_{n-1})$ such that $H1(x_1, \cdots, x_{n-1}) \neq K(x_1, \cdots, x_{n-1})$ we must have

$$(z_1, \cdots, z_{n-1}) < (v_1, \cdots, v_{n-1}). \tag{3.2}$$

Let $H2$ be obtained from $H1$ by putting $H2(v_1, \cdots, v_{n-1}) = K(v_1, \cdots, v_{n-1})$ and keeping $H2 = H1$ for all other function arguments. Then $H2$ will generate a de Bruijn sequence since the latter operation (changing $H1$ to $H2$) corresponds to joining $C_1$ and $C_2$. $H < H2$ as a consequence of (3.2), i.e. the de Bruijn sequence generated by $H2$ is obtained from the one generated by $H$ by the cross-join pair operation. This means that $H$ does not exist, since by definition it is impossible to obtain a de Bruijn sequence greater than the one generated by $H$ by the cross-join pair operation (applied to the one generated by $H$). □

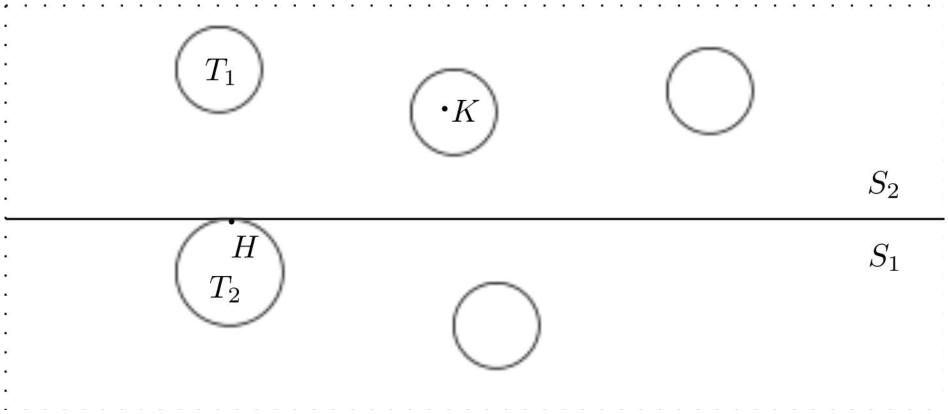

Fig. 1. Subsets and equivalence classes in the proof of Theorem 3.1.

$K, H, H1$ and $H2$ are functions mapping $\mathbb{F}_2^{n-1}$ to $\mathbb{F}_2$; $T_1, T_2$ are equivalence classes of such functions. All functions in $S_1$ are lexicographically smaller than those in $S_2$. We have $K \in S_2$, but it is irrelevant for the proof whether $K \in T_1$ or not. We have $K \notin T_2$ since $S_1 \cap S_2 = \emptyset$. Moreover $S_1 \cup S_2 = S$, the set of all functions generating de Bruijn sequences of order $n$. *A prori* there can be many equivalence classes of de Bruijn sequences, but we prove that there is only one.

Since there is a one-to-one correspondence between de Bruijn sequences of order $n$, which have length $2^n$ or period $2^n$ when considered as periodic sequences, and modified de Bruijn sequences of length (period) $2^n - 1$, Theorem 3.1 is literally true for modified de Bruijn sequences. As an immediate consequence of Theorem 3.1 we also have the following:

Lemma 3.2. *For every de Bruijn sequence (modified de Bruijn sequence) there exist cross-join pairs.*

According to Helleseth and Kløve [**11**] the number of different cross-join pairs for an $m$-sequence of order $n$ (modified de Bruijn sequence of maximum period $2^n - 1$ generated by the feedback function of *LFSR*) is equal to

$$(2^{n-1} - 1)(2^{n-1} - 2)/6. \tag{3.3}$$

For modified de Bruijn sequences of order $n$ generated by feedback functions of *NLFSRs* we do not have explicit formulae for the number of cross-join pairs, but our experiments show that for small orders $n = 4, 5, 6$ these numbers change with *NLFSRs* of given order and are around the value (3.3). *ANF* of the feedback function of a given de Bruijn sequence of order $n$ is different from that of the corresponding modified de Bruijn sequence of the same order (obtained by removing one zero from the tuple of $n$ consecutive zeros). If $f(x_0, x_1, \cdots, x_{n-1})$ is a Boolean feedback function which generates a modified de Bruijn sequence of order $n$, then the Boolean feedback function of the corresponding de Bruijn sequence is equal to

$$f(x_0, x_1, \cdots, x_{n-1}) + \overline{x}_1\overline{x}_2 \cdots \overline{x}_{n-1}. \tag{3.4}$$

According to (3.4) a Boolean feedback function which generates a de Bruijn sequence contains the term $x_1 \cdots x_{n-1}$; it has algebraic degree $n - 1$. Boolean feedback functions generating modified de Bruijn sequences of order $n$ have algebraic degree at most $n - 2$; some of them have degree 1 (those corresponding to *LFSRs* generating $m$-sequences) and there are some of low degrees $n = 2, 3$ or 4 and having a small number of terms in their *ANF*. In [**8**] Gammel *et al.* have found *NLFSRs* up to order of $n = 33$, algebraic degree up to 5 and containing up to around 30 terms in their *ANF*. According to Theorem 3.1 we know that such simple feedback functions of maximum period *NLFSRs* can be obtained by repeated application of the cross-join operation to a feedback function of *LFSR* generating an $m$-sequence. The problem is to find an effective algorithm to realize this task. The next section presents our efforts in this direction.

## 4. Applications

In this section we will consider modified de Bruijn sequences and feedback functions of *NLFSRs* generating these sequences. We use the cross-join pair operation to construct new *NLFSRs* from a given one. Let $f(x_0, x_1, \cdots, x_{n-1})$ be a feedback function generating a modified de Bruijn sequence. Let

$$\alpha = (a_0, a_1, \cdots, a_{n-1}), \quad \widehat{\alpha} = (\overline{a}_0, a_1, \cdots, a_{n-1}),$$

$$\beta = (b_0, b_1, \cdots, b_{n-1}), \quad \widehat{\beta} = (\overline{b}_0, b_1, \cdots, b_{n-1})$$

be a cross-join pair for the sequence generated by the function $f(x_0, x_1, \cdots, x_{n-1})$.

Then the function

$$f(x_0, x_1, \cdots, x_{n-1}) + (x_1 + a_1) \cdots (x_{n-1} + a_{n-1}) + (x_1 + b_1) \cdots (x_{n-1} + b_{n-1})$$

is a feedback of a new modified de Bruijn sequence. When we apply the cross-join operation several times to a maximum period *LFSR* it can happen that the resulting higher degree terms cancel and we obtain a feedback function with a simple *ANF*. In fact, we do not have strict control on the process of algebraic cancellation of terms resulting from cross-join pairs, and finding *NLFSRs* with a simple *ANF* is a random process.

**4.1. *NLFSRs* of order 4.** This example was a starting point for our investigations. We give here a list of feedback functions generating all modified de Bruijn sequences of order $n = 4$. Functions 1 and 2 represent primitive *LFSRs* of order 4. Each of the sequences generated by them has seven cross-join pairs. Applying Theorem 2.4 we get in total twelve new nonlinear feedback functions, since two of them appear twice (the red lines in Figure 2). But these are not all existing nonlinear feedback functions generating modified de Bruijn sequences. An application of two additional cross-join operations gives two missing nonlinear feedback functions (the edges (3,5) and (4,6) of the graph). One can see that any two vertices of the graph can be connected by a path which results as an application of several cross-join operations.

Now the problem arises whether one can implement an algorithm which takes as input a feedback function of a primitive *LFSR* and constructs by repeated application of the cross-join operation all feedback functions (linear and nonlinear) of modified de Bruijn sequences. We have implemented such an algorithm. It works for $n = 4$ (see Figure 3) and $n = 5$ giving all corresponding feedback functions. For $n = 5$ we have 2048 feedback functions altogether, see (2.2). The algorithm works for greater orders $n$ too. For $n = 6$ we have $2^{26}$ maximum period *NLFSRs*.

In our realization of the algorithm we have used the SAGE package [**22**], esspecially the *polybori* module implementing operations on algebraic normal forms of Boolean functions. All algorithm were implemented with the *Python* programming language [**21**].

Here is a list of all feedback functions (linear and nonlinear) generating 16 modified de Bruijn sequences of order 4. The *NLFSRs* are obtained by using the cross-join operation.

(1) $x_0 + x_1$
(2) $x_0 + x_3$
(3) $x_0 + x_1 + \overline{x}_1 x_2 x_3 + \overline{x}_1 x_2 \overline{x}_3 = x_0 + x_1 + x_2 + x_1 x_2$
(4) $x_0 + x_3 + \overline{x}_1 x_2 x_3 + \overline{x}_1 x_2 \overline{x}_3 = x_0 + x_2 + x_3 + x_1 x_2$
(5) $x_0 + x_1 + (\overline{x}_1 x_2 x_3 + \overline{x}_1 x_2 \overline{x}_3) + (x_1 x_2 \overline{x}_3 + x_1 \overline{x}_2 x_3)$
$= x_0 + x_1 + x_2 + x_1 x_3$
(6) $x_0 + x_3 + (\overline{x}_1 x_2 x_3 + \overline{x}_1 x_2 \overline{x}_3) + (x_1 x_2 \overline{x}_3 + x_1 \overline{x}_2 x_3)$
$= x_0 + x_2 + x_3 + x_1 x_3$
(7) $x_0 + x_3 + \overline{x}_1 x_2 \overline{x}_3 + \overline{x}_1 \overline{x}_2 x_3 = x_0 + x_2 + x_1 x_2 + x_1 x_3$
(8) $x_0 + x_1 + \overline{x}_1 x_2 \overline{x}_3 + \overline{x}_1 \overline{x}_2 x_3 = x_0 + x_1 + x_2 + x_3 + x_1 x_2 + x_1 x_3$
(9) $x_0 + x_1 + x_1 x_2 \overline{x}_3 + \overline{x}_1 x_2 \overline{x}_3 = x_0 + x_1 + x_2 + x_2 x_3$
(10) $x_0 + x_3 + x_1 x_2 \overline{x}_3 + \overline{x}_1 x_2 \overline{x}_3 = x_0 + x_2 + x_3 + x_2 x_3$
(11) $x_0 + x_1 + \overline{x}_1 x_2 x_3 + x_1 x_2 \overline{x}_2 = x_0 + x_1 + x_1 x_2 + x_2 x_3$
(12) $x_0 + x_1 + x_1 \overline{x}_2 \overline{x}_3 + \overline{x}_1 \overline{x}_2 x_3 = x_0 + x_3 + x_1 x_2 + x_2 x_3$
(13) $x_0 + x_1 + x_1 \overline{x}_2 \overline{x}_3 + \overline{x}_1 x_2 \overline{x}_3 = x_0 + x_2 + x_1 x_3 + x_2 x_3$
(14) $x_0 + x_3 + x_1 \overline{x}_2 \overline{x}_3 + x_1 \overline{x}_2 x_3 = x_0 + x_1 + x_2 + x_3 + x_1 x_3 + x_2 x_3$
(15) $x_0 + x_1 + x_1 \overline{x}_2 x_3 + \overline{x}_1 x_2 \overline{x}_3 = x_0 + x_1 + x_2 + x_1 x_2 + x_1 x_3 + x_2 x_3$
(16) $x_0 + x_3 + x_1 \overline{x}_2 x_3 + \overline{x}_1 x_2 \overline{x}_3 = x_0 + x_2 + x_3 + x_1 x_2 + x_1 x_3 + x_2 x_3$

We illustrate below the process of constructing *NLFSRs* of order 4 starting from two *LFSRs* (Fig. 2) and starting from one *LFSR* (Fig. 3).

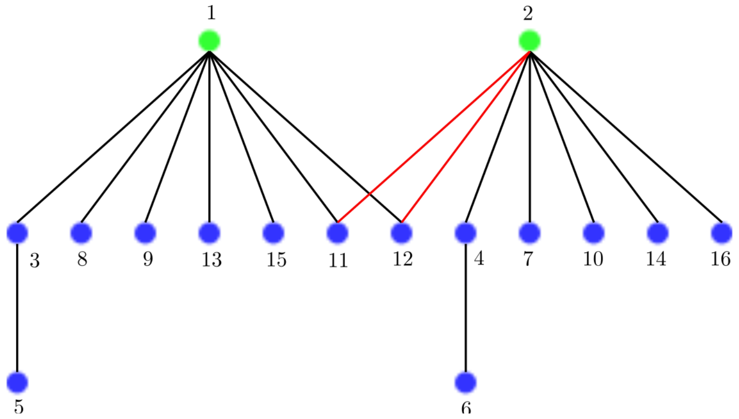

Fig. 2. The first graph of cross joining for *NLFSRs* of order 4.

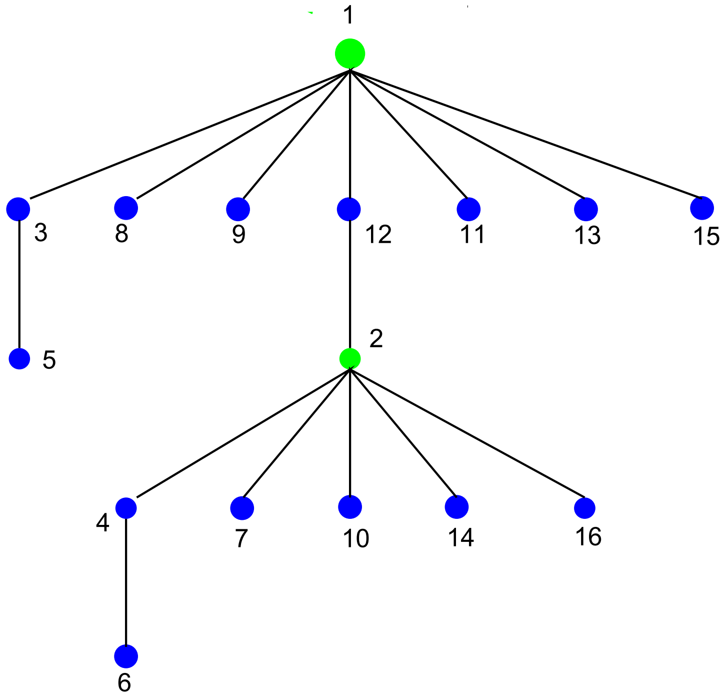

Fig. 3. The second graph of cross joining for *NLFSRs* of order 4.

**4.2. Quadratic $m$-sequences.** Chen, Games and Rushanan [**2**] have investigated the case when the feedback function (2.4) is a quadratic Boolean function of $n$ variables; i.e., has the following algebraic normal form:

$$f(x_0, x_1, \cdots, x_{n-1}) = \sum_{0\le i\le j\le n-1} a_{ij}x_ix_j. \tag{4.1}$$

Let us note that $x_i^2 = x_i$ for all $i \ge 0$, hence the coefficients $a_{ii}$ correspond to the linear terms of the function $f$. The recurrence (2.3) corresponding to the quadratic function (4.1) has the form

$$s_{n+k} = \sum_{0\le i\le j\le n-1} a_{ij}s_{i+k}s_{j+k} \tag{4.2}$$

for all $k \ge 0$. The authors of [**2**] have introduced a notion of quadratic $m$-sequences by analogy to the linear ones.

Definition 4.1. A binary sequence $\mathbf{s}$ is called a *quadratic $m$-sequence* of order $n$ (span $n$) if it satisfies the quadratic recurrence (4.2) and has period $2^n - 1$.

The authors of [**2**] have studied algorithmic generation of some quadratic $m$-sequences and listed them up to order $n = 12$. Namely, they considered quadratic Boolean functions of the form

$$f(x_0, x_1, \cdots, x_{n-1}) = g(x_0, x_1, \cdots, x_{n-1}) + x_i + x_ix_j, \tag{4.3}$$

where $i \ne j,\ 1 \le i, j \le n-1$ and

$$g(x_0, x_1, \cdots, x_{n-1}) = x_0 + c_1x_1 + \cdots + c_{n-1}x_{n-1}$$

is a linear function which generates an $m$-sequence, i.e., the corresponding polynomial in the ring $\mathbb{F}_2[x]$

$$p(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + 1$$

is primitive. Primitive polynomials of degree $n$ have their roots in the finite Galois field $\mathbb{F}_{2^n}$; these roots are primitive elements (generators) of the multiplicative group $\mathbb{F}_{2^n}^*$. It is known that there is a one-to-one correspondence between linear $m$-sequences of period $2^n - 1$ and primitive polynomials of degree $n$. The number of primitive polynomials of degree $n$ is equal to $\varphi(2^n-1)/n$, where $\varphi(.)$ is the Euler function. The proofs of all these facts can be found in the books [**9, 10, 13**].

The quadratic recurrences corresponding to Boolean functions (4.3) are modifications of the linear one. The term $x_i + x_i x_j$ introduces a nonlinear *perturbation* to the given $m$-sequence. The states of the *LFSR* for which the function $x_i + x_i x_j$ equals 1 break or join the corresponding cycles of the *LFSR*. The case when after running over all states of the *FSR* we get only one cycle is the sought-for *NLFSR*. In fact, this is a random phenomenon. The relevant discussion has been presented in [**2**]. In practice, not all primitive polynomials and terms $x_i + x_i x_j$ lead to a suitable coincidence giving a quadratic $m$-sequence.

In [**3**] we continued the search for those NLFSRs generating quadratic $m$-sequences up to order $n = 29$ and found the feedback Boolean function

$$f(x_0, \cdots, x_{28}) = x_0 + x_3 + x_5 + x_6 + x_{11} + x_{12} + x_{16} + x_{19} + x_{22} + x_{23} + x_{27} + x_{20}x_{28}$$

which generates a modified de Bruijn sequence of period $2^{29} - 1$. Here the term $x_{20}$ coming from a primitive polynomial of degree 29 cancels with such a term in the expression $x_{20} + x_{20}x_{28}$.

**4.3. A *NLFSR* of order 7.** We take the primitive polynomial $p(x) = x^7 + x + 1$. The feedback function of the corresponding LFSR is $x_0 + x_6$. We start from the initial state $s_0 = (x_0, x_1, \ldots, x_6) = (1, 0, \ldots, 0)$ and generate all nonzero states $s_1, \ldots, s_{126}$ of the LFSR. Let us consider the states for which $x_2 = 1$ and $x_4 = 0$; there are altogether 32 of them. We can find eight cross-join pairs having no common states, which all cover the states with $(x_2, x_4) = (1, 0)$. Here is a list of those cross-join pairs $(\alpha, \widehat{\alpha}; \beta, \widehat{\beta})$ :

$$(10, 29; 17, 101),\ (20, 58; 21, 91),\ (25, 47; 32, 62),\ (37, 118; 113, 125),$$

$$(38, 107; 71, 119),\ (42, 55; 50, 94),\ (59, 82; 81, 90),\ (65, 97; 70, 106).$$

We now apply the cross-join operation to each cross-join pair to obtain

$$\begin{aligned} f = x_0 + x_6 &+ x_1x_2x_3\overline{x}_4x_5\overline{x}_6 + \overline{x}_1x_2\overline{x}_3\overline{x}_4x_5x_6 + \overline{x}_1x_2x_3\overline{x}_4\overline{x}_5x_6 + x_1x_2\overline{x}_3\overline{x}_4x_5x_6 \\ &+ x_1x_2x_3\overline{x}_4x_5x_6 + \overline{x}_1x_2\overline{x}_3\overline{x}_4x_5\overline{x}_6 + \overline{x}_1x_2x_3\overline{x}_4\overline{x}_5\overline{x}_6 + \overline{x}_1x_2\overline{x}_3\overline{x}_4\overline{x}_5\overline{x}_6 \\ &+ x_1x_2\overline{x}_3\overline{x}_4\overline{x}_5x_6 + x_1x_2\overline{x}_3\overline{x}_4\overline{x}_5\overline{x}_6 + \overline{x}_1x_2x_3\overline{x}_4x_5x_6 + \overline{x}_1x_2x_3\overline{x}_4x_5\overline{x}_6 \\ &+ x_1x_2\overline{x}_3\overline{x}_4x_5\overline{x}_6 + x_1x_2x_3\overline{x}_4\overline{x}_5x_6 + \overline{x}_1x_2\overline{x}_3\overline{x}_4\overline{x}_5x_6 + x_1x_2x_3\overline{x}_4\overline{x}_5\overline{x}_6 \\ &= x_0 + x_6 + x_2 + x_2x_4. \end{aligned}$$

The last formula is a feedback function of NLFSR generating a modified de Bruijn sequence of period $2^7 - 1 = 127$.

**4.4. Necessary conditions.** Let us consider a candidate for a feedback function of *NLFSR* of maximum period $2^n - 1$ which has the form

$$f(x_0, x_1, \cdots, x_{n-1}) = g(x_0, x_1, \cdots, x_{n-1}) + h(x_1, \cdots, x_{n-1}), \tag{4.4}$$

where $g$ is a function as in (4.3) and $h$ is a Boolean function of low algebraic degree (say up to 5) and with the number of terms from 10 to 20. The number of all terms in $f$ must be even. In our construction we need to know the set $\mathcal{S}$ of states $(x_0, x_1, \cdots, x_{n-1})$ for which

$$h(x_1, \cdots, x_{n-1}) = 1,$$

and the number of states in $\mathcal{S}$ must be a multiple of 4. Now we find a set $\mathcal{C}$ of cross-join pairs for the $m$-sequence generated by the feedback function $g$, but only those whose states belong to $\mathcal{S}$. Suppose the collection $\mathcal{C}$ of cross-join pairs has the following properties:

(1) All cross-join pairs in $\mathcal{C}$ are disjoint; they do not contain common states from $\mathcal{S}$.
(2) All states in cross-join pairs of $\mathcal{C}$ exactly cover the set $\mathcal{S}$.

Then conditions 1 and 2 are necessary for the function (4.4) to generate a sequence of maximum period $2^n - 1$. An example of such a family of cross-join pairs is given above in Section 4.3. We have verified experimentally this construction of maximum period *NLFSRs* with feedback functions of the form (4.3) for orders up to $n = 17$. Conditions 1 and 2 are not sufficient to guarantee the maximum period. When they are satisfied one must check the period of the sequence generated by the feedback function $f$.

## References

[1] N. G. de Bruijn. *A combinatorial problem.* Indag. Math., 8(1946), pp. 461-467.

[2] A. H. Chan, R. A. Games, J. J. Rushanan. *On the quadratic m-sequences.* Proceedings of Fast Software Encryption. LNCS vol. 809, pp. 166-173. Springer-Verlag, 1994.

[3] P. Dabrowski, G. Labuzek, T. Rachwalik, J. Szmidt. *Searching for nonlinear feedback shift registers with parallel computing.* IACR Cryptology ePrint Archive 2013/542. 2013 Military Communications and Information Systems Conference. MCC 2013, Malto, France.

[4] E. Dubrova. *A scalable method for constructing Galois NLFSRs with period* $2^n - 1$ *using cross-join pairs.* IEEE Trans. on Inform. Theory, 59(1), 2013, pp. 703-709.

[5] C. Flye Sainte-Marie. *Solution to question nr. 48.* L'Intermédiaire des Mathématiciens, 1(1894). pp. 107-110.

[6] H. Fredricksen. *A class of nonlinear de Bruijn cycles.* J. of Combinatorial Theory (A), 19(1975), pp. 192-199.

[7] H. Fredricksen. *A survey of full length nonlinear shift register cycle algorithms.* SIAM Review, 24(2), 1982, pp. 195-221.

[8] B. M. Gammel, R. Goetffert, O. Kniffler. *Achterbahn 128/80.* The eSTREAM project, www.ecrypt.eu.org/stream/, www.matpack.de/achterbahn/

[9] S. W. Golomb. *Shift register sequences.* San Francisco, Holden-Day, 1967, revised edition, Laguna Hills, CA, Aegean Park Press, 1982.

[10] S. W. Golomb, G. Gong. *Signal Design for Good Correlation. For Wireless Communication, Cryptography, and Radar.* Cambridge University Press, 2005.

[11] T. Helleseth, T. Kløve. *The number of cross-join pairs in maximum length linear sequences.* IEEE Trans. on Inform. Theory, 31(1991), pp. 1731-1733.

[12] C. J. A. Jansen. *Investigations on nonlinear streamcipher systems: Construction and evaluation methods.* Ph.D. Thesis, Technical University of Delft, 1989.

[13] R. Lidl, H. Niederreiter. *Introduction to Finite Fields and their Applications (Revised Edition).* Cambridge University Press, Cambridge, 1994.

[14] K. B. Magleby. *The synthesis of nonlinear feedback shift registers.* Technical Report no. 6207-1. Stanford Electronics Laboratories, 1963.

[15] K. Mandal, G. Gong. *Cryptographically strong de Bruijn sequences with large periods.* Selected Areas in Cryptography. L. R. Knudsen, K. Wu (Eds.). LNCS, vol. 7707, pp. 104-118. Springer-Verlag, 2012.

[16] G. L. Mayhew, S. W. Golomb. *Linear spans of modified de Bruijn sequences.* IEEE Trans. Inform. Theory, 36(5), 1990, pp. 1166-1167.

[17] J. Mykkeltveit. *Generating and counting the double adjacencies in a pure cycing shift register.* Trans. on Computers, C-24, 1975, pp. 299-304.

[18] J. Mykkeltveit, M-K. Siu, P. Tong. *On the cyclic structure of some nonlinear shift register sequences.* Inform. and Control, 43(1979), pp. 202-215.

[19] T. Rachwalik, J. Szmidt, R. Wicik, J. Zabocki. *Generation of nonlinear feedback shift registers with special purpose hardware.* 2012 Military Communications and Information Systems Conference. MCC 2012, pp. 151-154. IEEE Xplore Digital Library. IACR Cryptology ePrint Archive 2012/314.

[20] M. S. Turan. *On the nonlinearity of maximum-length NFSR feedbacks.* Cryptography and Communications, 4(3-4), 2012, pp. 233-243.

[21] *Python Programming Language.* http://www.python.org

[22] *SAGE Mathematical Software.* Version 5.8. http://www.sagemath.org

International Research Institute of Stavanger (IRIS), Bergen, Norway

Miltary Communication Institute, ul. Warszawska 22A, 05-130 Zegrze, Poland
*E-mail address*: j.szmidt@wil.waw.pl