Generating Good Span *n* Sequences Using Orthogonal Functions in Nonlinear Feedback Shift Registers

Kalikinkar Mandal and Guang Gong

Abstract A binary span n sequence generated by an n-stage nonlinear feedback shift register (NLFSR) is in a one-to-one correspondence with a de Bruijn sequence and has the following randomness properties: period $2^n - 1$, balance, and ideal n-tuple distribution. A span n sequence may have a high linear span. However, how to find a nonlinear feedback function that generates such a sequence constitutes a long-standing challenging problem for about 5 decades since Golomb's pioneering book, *Shift Register Sequences*, published in the middle of the 1960s. In hopes of finding good span n sequences with large linear span, in this chapter we study the generation of span n sequences using orthogonal functions in polynomial representation as nonlinear feedback in a nonlinear feedback shift register. Our empirical study shows that the success probability of obtaining a span n sequence in this technique is better than that of obtaining a span n sequence in a random span n sequence generation method. Moreover, we analyze the linear span of new span n sequences, and the linear span of a new sequence lies between $2^n - 2 - 3n$ (near optimal) and $2^n - 2$ (optimal). Two conjectures on the linear span of new sequences are presented and are valid for $n \leq 20$.

1 Introduction

Nonlinear feedback shift registers (NLFSRs) are used to design many cryptographic primitives such as pseudorandom sequence generators (PRSGs), stream ciphers [11], and lightweight block ciphers [7] for providing security and privacy in communication systems. Ciphers based on NLFSRs are of great practical importance in many constrained environments, for instance, RFID tags and sensor networks due to their need for efficient hardware implementation and high throughput. In general, an arbitrary NLFSR cannot be used for generating keystreams in stream ciphers

K. Mandal (✉) • G. Gong
Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1
e-mail: kmandal@uwaterloo.ca; ggong@uwaterloo.ca

© Springer International Publishing Switzerland 2014

Ç.K. Koç (ed.), *Open Problems in Mathematics and Computational Science*,
DOI 10.1007/978-3-319-10683-0_7

because the randomness properties including the period of a sequence generated by that NLFSR are unknown and hard to determine.

A binary *de Bruijn sequence* is a binary sequence with period 2^n which satisfies the property that all n-tuples occur exactly once in one period. De Bruijn sequences have known randomness properties, namely, maximum period, balance property, ideal n-tuple distribution, and large linear span [3, 21, 34]. A *modified de Bruijn sequence* or *span n sequence* with period $2^n - 1$ is a pseudorandom sequence where each nonzero n-tuple occurs exactly once in one period of the sequence. This property is referred to as *the span n property* [21]. Often, de Bruijn sequences as well as span n sequences are generated recursively by an n-stage nonlinear feedback shift register. Only, m-sequences are a class of span n sequences generated by linear feedback shift registers.

A span n sequence can be constructed from a de Bruijn sequence by removing any one zero from the run of zeros of length n, and similarly, a de Bruijn sequence can be formed from a span n sequence by adding one zero to the run of zeros of length $n - 1$. The *linear span or linear complexity* of a sequence is the length of the shortest LFSR that produces the given sequence. We remember that "linear span" and "span n" are two different properties of a span n sequence. Note that by adding an extra zero to the run of zeros of length $n - 1$ to an m-sequence, the linear span of the resultant de Bruijn sequence varies between $2^{n-1} + n$ and $2^n - 1$ [3], but by removing any one zero from the run of zeros of length n from the resultant de Bruijn sequence, it becomes an m-sequence or a span n sequence with linear complexity n. So the lower bound of the linear span of the span n sequence drops to n [23]. This phenomenon suggests to study the randomness properties, particularly, the linear span property of span n sequences instead of de Bruijn sequences for cryptographic usages. Until recently, there is no known general construction of a nonlinear feedback function which generates a span n sequence, and this is open since the last 5 decades. Therefore, the generation of span n sequences by NLFSRs is a challenging problem.

Our objective is to produce span n/de Bruijn sequences using orthogonal functions as feedback functions in nonlinear feedback shift registers. An orthogonal feedback function has a trace representation and is composed of three parameters, namely, a decimation number, a primitive polynomial, and a t-tap position ($5 \leq t \leq n - 1$). In an NLFSR, a class of feedback functions is constituted by varying the decimation numbers and the polynomial bases of the finite fields. Finding span n sequences by using this class of feedback functions and all possible tap positions of the feedback functions is called a *structured search*. We show that a number of new span n sequences with a moderate n can be produced through the structured search. For $n \geq 10$, all the feedback functions of degree greater than or equal to two cannot be employed to search span n sequences. Using the structure search, on the other hand, one can employ a number of feedback functions with different degrees and a variable number of terms.

In this chapter, we present some new theoretical results on generating span n sequences and experimental results on finding the number of new span n sequences. The chapter is organized as follows. In Sect. 2, we provide some basic definitions

of shift register sequences and their properties. Section 2.2 recalls the definitions of known orthogonal functions, and Sect. 3 introduces some known constructions of de Bruijn sequences. In Sect. 4, we describe the span n sequence generation technique using orthogonal functions and develop some properties of this technique, including an estimation of the number of orthogonal feedback functions used in this technique. Sections 5 and 6 present the experimental results on the number of span n sequences produced using orthogonal functions, and Sect. 7 presents an empirical success probability comparison of obtaining span n sequences using orthogonal functions. In Sect. 8, we analyze the linear span of newly produced span n sequences by the aforementioned orthogonal functions and present two conjectures on the linear span of the span n sequences produced by the orthogonal functions. Our empirical results show that the success probability of obtaining a span n sequence in the structured search is larger than that of generating a span n sequence in a random search. Our results show that the linear span of a new span n sequence lies in the range of $2^n - 2 - 3n$ (near optimal) and $2^n - 2$ (optimal). In Sect. 9, some applications of new span n sequences are shown, and in the section "Conclusions", we conclude the chapter.

2 Preliminaries

In this section, we define and explain the terms and mathematical functions that will be used in this chapter to produce span n sequences.

- $\mathbb{F}_2 = \{0, 1\}$: the Galois field with two elements.
- $\mathbb{F}_{2^t} = \{(x_0, x_1, \ldots, x_{t-1}) \mid x_i \in \mathbb{F}_2\}$—an extension field that is defined by a primitive element α with $p(\alpha) = 0$, where $p(x) = c_0 + c_1 x + \cdots + c_{t-1}x^{t-1} + x^t$ is a primitive polynomial of degree t (≥ 2) over $\mathbb{F}_2$.
- $\mathrm{Tr}(x) = x + x^2 + \cdots + x^{2^{t-1}}$: the trace function mapping from $\mathbb{F}_{2^t}$ to $\mathbb{F}_2$.
- $D_t = \{d : d$ is a coset leader with $\gcd(d, 2^t - 1) = 1\}$. The cardinality of D_t, denoted as $|D_t|$, is given by $\frac{\phi(2^t-1)}{t}$, where $\phi(\cdot)$ is the Euler phi function.

2.1 Basic Definitions and Properties of Feedback Shift Registers

Usually, an n-stage linear or nonlinear feedback shift register is used to generate a periodic binary sequence $\mathbf{a} = \{a_i\}$, and the recurrence relation for the (N)LFSR is defined as [20]

$$a_{n+k} = a_k \oplus g(a_{k+1}, \ldots, a_{k+n-1}),\ a_i \in \mathbb{F}_2,\ k \geq 0$$

where $(a_0, a_1, \ldots, a_{n-1})$ is the *initial state* of the shift register, g is a Boolean function in $(n-1)$ variables, and $\oplus$ is the addition operation over $\mathbb{F}_2$. If the function g is an affine function, then the sequence $\mathbf{a}$ is called an LFSR sequence; otherwise, it is called an NLFSR sequence. The above recurrence relation is also known as a nonsingular recurrence relation.

The complementary binary sequence of binary sequence $\mathbf{b} = \{b_i\}_{i\geq 0}$, denoted as $\bar{\mathbf{b}}$, is defined by $\{\bar{b_i}\}_{i\geq 0}$, where $\bar{b_i} = b_i \oplus 1$. The *linear span or linear complexity* of a sequence is the length of the shortest LFSR that produces the sequence.

Definition 1 ([22]) The autocorrelation of a binary sequence $\{a_i\}$ with period N is defined as

$$C(\tau) = \sum_{i=0}^{N-1} (-1)^{a_{i+\tau}+a_i}.$$

Moreover, if $N = 2^n - 1$, the sequence $\{a_i\}$ has 2-level autocorrelation if

$$C(\tau) = \begin{cases} 2^n - 1 & \text{if } \tau \equiv 0 \pmod{2^n - 1} \\ -1 & \text{if } \tau \not\equiv 0 \pmod{2^n - 1}. \end{cases}$$

Property 1 The linear span of a de Bruijn sequence, denoted as LS_{db}, is bounded by [3]

$$2^{n-1} + n \leq LS_{\text{db}} \leq 2^n - 1. \tag{1}$$

On the other hand, the linear span of a span n sequence that is generated by an NLFSR, denoted as LS_s, is bounded by

$$2n < LS_s \leq 2^n - 2. \tag{2}$$

From this property, we say that a span n sequence has the optimal linear span if its linear span is equal to $2^n - 2$.

2.2 Review of the Trace Representation of 2-level Autocorrelation Sequences

An orthogonal function from $\mathbb{F}_{2^t}$ to $\mathbb{F}_2$ is in one-to-one correspondence with a binary sequence with (ideal) 2-level autocorrelation function, 2-level autocorrelation sequence in short. There are only very few known constructions on 2-level autocorrelation sequences, which constitutes another challenge problem for years. Interestingly, those functions possess good cryptographic properties. The reader is referred to Golomb and Gong's book [22] for the details about the constructions

of 2-level autocorrelation sequences and their related cryptographic properties. In the following, for easy reference, we formally provide the definition of orthogonal functions and their known constructions from corresponding trace representation of 2-level autocorrelation sequences.

Definition 2 A function, say, $f(x)$, from $\mathbb{F}_{2^t}$ to $\mathbb{F}_2$ is called an orthogonal function if $\sum_{x\in\mathbb{F}_{2^t}}(-1)^{f(\lambda x)+f(x)} = 0$ for all $(1 \neq)\lambda \in \mathbb{F}_{2^t}$.

Let α be a primitive element of $\mathbb{F}_{2^t}$ and let $a_i = f(\alpha^i)$ where the binary sequence $\{a_i\}$ is called an evaluation of $f(x)$ and $f(x)$, the trace representation of $\{a_i\}$.

Property 2 With the above notation:

1. $f(x)$ is orthogonal if and only if its evaluation has 2-level autocorrelation.
2. If $f(x)$ is orthogonal, then $f(x^r)$ is orthogonal for all r with $(r, 2^t - 1) = 1$.

Let $C = \{r, 2r, \ldots, 2^{n_r-1}r\}$ where n_r is the smallest number such that $r2^{n_r} \equiv r \bmod 2^t - 1$. Then C is called a (cyclotomic) coset consisting r modulo $2^t - 1$, and the smallest number in C is called the coset leaders of C. Let I consist of all coset leaders modulo $2^t - 1$.

2.2.1 Number Theory-Based Constructions

This type of the constructions includes Legendre sequences and Hall sextic residue sequences. Let $p = 2^t - 1$ be a prime number, u be a primitive element in $\mathbb{F}_p$, and $c = \frac{2^t-2}{t}$.

Orthogonal Functions from Legendre Sequences (A1) Let

$$f(x) = \sum_{i=0, i\in I}^{c/2-1} \mathrm{Tr}(x^{u^{2i}}), x \in \mathbb{F}_{2^t}.$$

Or equivalently,

$$f(x) = \sum_{i\in I_0} \mathrm{Tr}(x^i), x \in \mathbb{F}_{2^t}$$

where $I_0 \subset I$ consist of all quadratic coset leaders modulo $2^t - 1$. Then $f(x)$ is an orthogonal function from $\mathbb{F}_{2^t}$ to $\mathbb{F}_2$ whose evaluation gives a Legendre sequence with 2-level autocorrelation.

Hall's "Sextic Residue Sequence" (A2) Additional to the Legendre sequences, $p = 4t - 1 = 4a^2 + 27$. Let

$$f(x) = \sum_{i=0, i\in I}^{c/6-1} \mathrm{Tr}(x^{u^{6i}}), x \in \mathbb{F}_{2^t}.$$

Then $f(x)$ is an orthogonal function from $\mathbb{F}_{2^t}$ to $\mathbb{F}_2$ whose evaluation gives a Hall's "sextic residue sequence" with 2-level autocorrelation function.

2.2.2 Finite Fields-Based Constructions

There are four types of constructions for 2-level autocorrelation sequences: m-sequences, hyperoval constructions, Welch–Gong transformation construction, and Kasami power function construction including three-term and five-term sequences.

Orthogonal Functions from m-Sequences Let

$$f(x) = \mathrm{Tr}(x), x \in \mathbb{F}_{2^t},$$

then $f(x)$ is an orthogonal function whose evaluation gives an m-sequence with period $2^t - 1$, and the other m-sequences are given by $\mathrm{Tr}(x^d)$ where $\gcd(d, 2^t - 1) = 1$.

Orthogonal Functions from Hyperoval Sequences There are three monomial hyperoval sequences with 2-level autocorrelation, namely, Segre type and Glynn type 1 and type 2. Except for Segre hyperoval sequences, the trace representation is not represented in a formula. Instead, it is described in terms of some relation which needs to be computed for different t.

Let $(1)^l$ denote a string of l consecutive 1s. Let $\mathcal{A}$ denote the set consisting of all strings of the form $(1)^{4a+1}0$ or $a \geq 0$ and $(1)^{4b}$, $b \geq 0$. Let $\mathcal{A}^*$ denote the set of all strings obtained by concatenating zero, one or more strings from $\mathcal{A}$. Let t be a prime and

$$\begin{array}{|l|}\hline 01(\text{string in } \mathcal{A}^*)0(1)^{2s}, s \geq 0 \text{ or} \\ \hline 011(\text{string in } \mathcal{A}^*)11 \\ \hline \end{array}. \tag{3}$$

The trace representation of a Segre hyperoval sequence of period $2^t - 1$, t odd, is given by

$$f(x) = \sum_{i \in T_{\text{Segre}}} \mathrm{Tr}(x^i), x \in \mathbb{F}_{2^t}$$

where $T_{\text{Segre}} \subset I$ which are the collections of coset leaders of the all binary numbers given by (3) [5].

Let $T_{\text{Glynn}} \subset I$ be the collections of the coset leaders of solutions to

$$w(j) + w((k-1)j - w(kj) = 1, j = 1, \ldots, 2^t - 1$$

where $w(x)$ is the Hamming weight of binary number x. Then the trace representation of a Glynn hyperoval sequence of period $2^t - 1$, t odd, is given by

$$f(x) = \sum_{i \in T_{\text{Glynn}}} \text{Tr}(x^i), x \in \mathbb{F}_{2^t}$$

where $k = \sigma + \gamma$ for Glynn type 1 and $k = 3\sigma + 4$ for Glynn type 2 where $\sigma = 2^{(t+1)/2}$ and $\gamma = 2^{(3t+1)/4}$ [13].

Orthogonal Functions from Three-Term, Five-Term, and Welch–Gong Transformation Constructions In [38], it was conjectured that three-term and five-term sequences have 2-level autocorrelation as well as Welch–Gong transformation sequences discovered by Golomb, Gong, and Gaal. The validity of those conjectures is established later on by Dillon and Dobbertin in [8, 9].

Let $t = 2k - 1$ for some positive integer k and $t \geq 5$. Let

$$f(x) = \text{Tr}(x + x^{2^k+1} + x^{2^k-1}), x \in \mathbb{F}_{2^t}.$$

Then its evaluation gives three-term 2-level autocorrelation sequences.

Let t be a positive integer with $t \bmod 3 \not\equiv 0$ and $3k \equiv 1 \bmod t$ for some integer k. We define the function h from $\mathbb{F}_{2^t}$ to $\mathbb{F}_{2^t}$ by

$$h(x) = x + x^{q_1} + x^{q_2} + x^{q_3} + x^{q_4}$$

where

$$q_1 = 2^k + 1, q_2 = 2^{2k} + 2^k + 1, q_3 = 2^{2k} - 2^k + 1, q_4 = 2^{2k} + 2^k - 1.$$

(Note that $h(x)$ is a permutation over $\mathbb{F}_{2^t}$ [8].) Let

$$g(x) = \text{Tr}(h(x)) \text{ and } f(x) = \text{Tr}(h(x+1) + 1)$$

where $f(x)$ is known as the *WG transformation*. The evaluations of $g(x)$ and $f(x)$ yield five-term sequences and WG transformation sequences.

Orthogonal Functions from Kasami Power Function Construction Let $\gcd(k, t) = 1, k < t, kk' \equiv 1$, and

$$f(x) = \text{Tr}(R(x)), x \in \mathbb{F}_{2^t}$$

where $R(x)$ is given by

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

where A_i and V_i are iteratively defined by

$$\begin{aligned} A_1(x) &= x \\ A_2(x) &= x^{2^k+1} \\ A_{i+2}(x) &= x^{2^{(i+1)k}} A_{i+1}(x) + x^{2^{(i+1)k}-2^{ik}} A_i(x), i \geq 1 \end{aligned}$$

and

$$\begin{aligned} V_1(x) &= 0 \\ V_2(x) &= x^{2^k-1} \\ V_{i+2}(x) &= x^{2^{(i+1)k}} V_{i+1}(x) + x^{2^{(i+1)k}-2^{ik}} V_i(x), i \geq 1. \end{aligned}$$

Orthogonal Functions from Subfield Constructions Let $1 < m \mid t$, $m \neq t$, and $g(x)$ be any orthogonal function from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$, listed in the above subsections, and let

$$f(x) = \mathrm{Tr}_m^t(g(x)), x \in \mathbb{F}_{2^t}$$

where $\mathrm{Tr}_m^t(x)$ is the trace function from $\mathbb{F}_{2^t}$ to $\mathbb{F}_{2^m}$, i.e.,

$$\mathrm{Tr}_m^t(x) = x + x^{2^m} + \cdots + x^{2^{(l-1)m}}, x \in \mathbb{F}_{2^t}, l = t/m.$$

Then $f(x)$ is an orthogonal function from $\mathbb{F}_{2^t}$ to $\mathbb{F}_2$, and its evaluation is called a subfield 2-level autocorrelation sequences which includes GMW sequence for $g(x) = \mathrm{Tr}_1^m(x^d)$ where $\mathrm{Tr}(x)$ is the trace function from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$ and $\gcd(d, 2^m - 1) = 1$ and generalized GMW sequences for the rest of $g(x)$. Here we shorten them as GMW sequences.

2.2.3 Orthogonal Functions for Small Fields

In the following, we give the exponents explicitly for all known orthogonal functions of the form $f(x) = \sum_{i \in I} \mathrm{Tr}(x^i)$ from $\mathbb{F}_{2^t}$ to $\mathbb{F}_2$ for $5 \leq t \leq 11$ in Tables 1, 2 and 3 where the monomial function $\mathrm{Tr}(x)$ is not listed.

Table 1 Exponents in the orthogonal functions over $\mathbb{F}_{2^t}$, $5 \leq t \leq 9$

Orthogonal functions	Trace spectra	# of terms
$t = 5$		
T3	1, 3, 5	3
$t = 7$		
T3	1, 9, 13	3
T5	1, 5, 21, 13, 29	5
WG	1, 3, 7, 19, 29	5
QR	3, 5, 7, 23, 27, 29, 43, 55, 63	9
Hall	5, 27, 63	3
$t = 8$		
GMW	7, 13, 37, 11	4
T5	1, 9, 37, 29, 39	5
WG	13, 19, 21, 29, 39	5
$t = 9$		
T3	1, 17, 25	3
GMW	3, 17, 129	3
Segre	1,5,7, 9, 19, 25, 37, 77, 117	9
Glynn 1	1, 5, 9, 13, 19, 37, 43	7
Glynn 2	17, 23, 37, 43, 45, 75, 87	7

Table 2 Exponents in the orthogonal functions over $\mathbb{F}_{2^t}$, $t = 10$

Orthogonal functions	Trace spectra	# of terms
T5	1, 9, 57, 73, 121	5
WG	1, 3, 5, 7, 11, 13, 15, 35, 69, 71, 89, 105, 121	13
GMW1	3, 17	2
GMW2	5, 9	2
GMW3	7, 19, 25, 69	4
GMW4	11, 13, 21, 73	4
GMW5	1, 5, 7, 9, 19, 25, 69	7
GMW6	15, 23, 27, 29, 77, 85, 89, 147	8
GMW7	3, 7, 11, 13, 15, 21, 23, 27, 29, 73, 77, 85, 89, 147	14

We define the following set:

$$D_t^* = \{d \ : \ d \in D_t \text{ and } f_d(\cdot), \text{ is nonlinear and } f_d(x) \neq f_{d_1}(x), d \neq d_1 (\in D_t^*)\}.$$

For all decimation numbers in D_t^*, we take into account all distinct orthogonal functions obtained from an orthogonal function using decimations.

Table 3 Exponents in the orthogonal functions over $\mathbb{F}_{2^t}, t = 11$

Orthogonal functions	Trace spectra	# of terms
T3	1, 33, 49	3
T5	1, 17, 121, 137, 143	5
WG	21, 23, 29, 35, 37, 41, 71, 89, 139, 165, 213, 307, 415	13
Segre = B2	1,5, 13, 21, 53, 77, 85, 205, 213, 309, 333, 341, 413, 423, 469	15
Glynn 1	1, 5, 9, 13, 19, 37, 43, 67, 69, 137, 163, 211, 293	13
Glynn 2	1, 5, 13, 17, 29, 37, 49, 61, 69, 81, 93, 101, 113, 125, 139, 147, 151, 157, 171, 173, 183	21
B3	1, 5, 7, 9, 19, 25, 81, 169, 295	9

3 Review of Known Constructions of (Modified) de Bruijn Sequences

There is a one-to-one correspondence between a de Bruijn sequence and a modified de Bruijn/span n sequence. When the construction of a feedback function that generates a span n sequence is known, the construction of a de Bruijn sequence can be known and vice versa. In this section, we provide some known de Bruijn and span n sequence generation techniques.

3.1 Known Constructions for de Bruijn Sequences

Problem of generating a de Bruijn sequence is easy to understand, but providing a solution for generating a de Bruijn sequence efficiently is a challenging problem. This problem is studied from algorithmic, graph theoretic, and algebraic technique points of view in the literature. In particular, generating a de Bruijn sequence using a feedback shift register is an algebraic technique, which exploits properties of a feedback function. In the following, we present some well-known approaches of constructing de Bruijn sequences.

3.1.1 Lempel's D-Morphism-Based Techniques for de Bruijn Sequences

Lempel in [26] proposed the concept of generating a de Bruijn sequence of period 2^{n+1} by first computing two D-morphic preimages of a de Bruijn sequence of period 2^n and then concatenating these two preimages at a conjugate pair. In this construction, it is assumed that the construction of the de Bruijn sequence of period 2^n is known. Later on, Annexstein in [1] and Chang et al. in [6] proposed two algorithms based on Lempel's D-homomorphism for producing de Bruijn sequences of long period. Games [18] proposed a generalized construction of

Lempel's construction in which a de Bruijn sequence of period 2^{n+1} is constructed from two different de Bruijn sequences of period 2^n using Lempel's conjugate.

In [36], Mykkeltveit et al. presented Lempel's construction in the form of a composited recurrence relation. Following Mykkeltveit et al.'s construction, Mandal and Gong in [28] refined and studied the composited construction, for producing strong composited de Bruijn sequences of arbitrarily long period from a span n sequence. For the properties and cycle structures of composited recurrence relations, see [27, 36]. Note that, in the composited construction, the feedback function of a de Bruijn sequence is a bit complicated, which contains a number of sum-of-product terms. Recently, Mandal and Gong in [29] analyzed composited de Bruijn sequences from D-morphic point of view and presented an iterative technique for computing the nonlinear feedback function of a composited de Bruijn sequence. In the composited construction one needs to know the construction of a feedback function of a span n sequence in order to generate a de Bruijn sequence of long period.

3.1.2 Algorithms for de Bruijn Sequence Generation

Fredricksen and Kessler in [16] proposed an algorithm based on lexicographic compositions for constructing de Bruijn sequences of period 2^n, and the amount of storage required in implementing the algorithm is linear in n. Fredricksen and Maiorana in [17] presented an algorithm for generating necklaces of length n in k colors, and a k-ary de Bruijn sequence of period k^n is produced by juxtaposing in order the periodic reductions of the necklaces.

Fredricksen [14] developed an algorithm to generate nonlinear de Bruijn sequences, and the algorithm requires $3n$ units of storage and outputs one bit in around n units of time. Fredricksen also exhibited that new de Bruijn sequences can be obtained from a de Bruijn sequence by cross-joining, and the number of such new de Bruijn sequences is 2^{2n-5}. The storage requirement for implementing the method is about $6n$ units. When this method is compared with Mandal and Gong's iterative technique (MG iterative technique) for composited de Bruijn sequences, MG iterative technique for the composited feedback function requires less amount of time as well as memory.

Etzion and Lempel [12] developed a construction of de Bruijn sequences with linear complexity $(2^{n-1} + n)$ for all $n \geq 3$. A detailed survey by Fredricksen of many other de Bruijn sequence generation techniques can be found in [15].

3.1.3 Cycle Joining Techniques for de Bruijn Sequence Generation

Cycle joining technique is one of the well-known methods of generating a de Bruijn sequence in which a de Bruijn sequence is constructed by joining a finite number of cycles produced by a feedback shift register. In this technique, first a feedback

function of a nonsingular feedback shift register is chosen, and then a different feedback function for a de Bruijn sequence is constructed from the first feedback function based on its cycle decomposition.

Jansen et al. [25] presented a cycle joining algorithm for generating de Bruijn sequences where the feedback function of a de Bruijn sequence is the sum of two functions; one function is the feedback function itself, and another function is constructed from the feedback function for joining cycles. In [25], it is shown that $O(2^{\frac{2n}{\log(2n)}})$ de Bruijn sequences of period 2^n can be produced when all irreducible polynomials of degree n is taken in a feedback shift register. The storage requirement for this method is $3n$ bits, and $4n$-unit of time is required to generate each bit of a de Bruijn sequence. A storage-time comparison between this algorithm and the MG iterative technique can be found in [6].

Yang and Dai in [40] proposed a construction of an m-ary de Bruijn sequence based on joining the cycles using modification sets of a feedback function f. In the construction, a nonlinear feedback function F of a de Bruijn sequence is constructed from the feedback function f using the modification sets of f. The authors showed that, when a circulating register is chosen, at least $2^{(\frac{m^n}{n}-mn)}$ feedback functions that generate de Bruijn sequences can be constructed. However, this method is not efficient for large values of n, since the method requires the cycles decomposition of f to construct the function F, and for a large n, it is very hard to obtain the cycle decomposition of f. Moreover, the feedback function would contain many product terms for joining of the cycles.

Hauge and Helleseth [24] proposed a technique based on an irreducible polynomial and its adjacency graph to generate de Bruijn sequences. In this technique, a de Bruijn sequence is obtained as maximum spanning trees from the adjacency graph of a feedback function corresponding to an irreducible polynomial. The lower bound for the number of de Bruijn sequences is determined in terms of the cyclotomic numbers.

3.2 *Known Techniques for Generating Modified de Bruijn Sequences*

Most of the research efforts devoted on span n sequences have been concerned about the number of span n sequences and the characteristics of nonlinear feedback functions [21, 33, 34] including the number of terms in the feedback functions [33, 35] and the weight of truth tables of the feedback functions [32, 33]. Mayhew and Golomb reported the number of span n sequences for different values of the linear span of span n sequences and for different values of the number of terms in the feedback functions ($4 \leq n \leq 6$) [34, 35]. Mayhew reported the number of span n sequences for different weight classes of the truth tables of the feedback functions for $n = 6$ [33]. However, the task of finding the number of span n sequences for

different weight classes and for different values of the linear span is an unsolved problem for $n \geq 7$.

In [4], Chan et al. have considered the generation of quadratic m-sequence that uses very simple quadratic functions as the feedback function, which is the sum of a linear function in n variables and a quadratic term for any two variables and reported the number of span n sequences for $5 \leq n \leq 12$. Dubrova in [10] and Rachwalik et al. in [39] found a few quadratic m-sequences, i.e., span n sequence generated using quadratic feedback functions for $4 \leq n \leq 24$ and $25 \leq n \leq 27$, respectively. Gammel et al. have searched span n sequences while designing stream cipher Achterban:128/80 based on nonlinear feedback shift registers [19].

Note that the feedback functions of an NLFSR in [10, 19, 39] contain only a few terms and are of low algebraic degree. All the methods for finding the number of span n sequences and verifying the span n property of a sequence use an exhaustive search method which is an exponential time algorithm in n.

4 A New Construction

In this section we first describe the recurrence relation of nonlinear feedback shift registers whose feedback functions are orthogonal functions. In an n-stage NLFSR, the feedback function can also be regarded as a Boolean function in t variables where $5 < t \leq n-1$. Our considered orthogonal feedback functions in t variables are balanced as the evaluation of the feedback function has 2-level autocorrelation and have even Hamming weight 2^{t-1}. Thus, the new span n sequences generated by a class of feedback functions belong to the weight class 2^{n-2}. Then we calculate the approximate number of feedback functions used in the structured search.

4.1 Description of Span n Sequence Generation Using Orthogonal Function

Let $\mathbf{a} = \{a_i\}$ be a binary sequence generated by an n-stage NLFSR whose nonlinear recurrence relation is defined as

$$a_{n+k} = a_k \oplus f_d(x_k),\ x_k = (a_{r_1+k}, a_{r_2+k}, \ldots, a_{r_t+k}) \in \mathbb{F}_{2^t}, d \in D_t^*,$$
$$0 < t < n,\ k \geq 0 \tag{4}$$

where $(r_1, r_2, \ldots, r_t)$ with $0 < r_1 < r_2 < \cdots < r_t \leq n-1$ is called a *t-tap position* of the NLFSR, $f_d(x) = f(x^d)$, $f(x)$ is an orthogonal function, and $\oplus$ is the addition over $\mathbb{F}_2$. For a proper selection of a t-tap position and a feedback function $f_d(x)$, the binary sequence $\mathbf{a}$ can be a *span n sequence*. We note that for any choice of a t-tap position and a feedback function $f_d(x)$, the binary sequence

may not be a span n sequence. The reason for choosing $t \leq (n-1)$ is to involve a small number of state variables in the feedback functions, which is benefited to the implementation of the NLFSR as well as the production of more feedback functions.

Let $\mathbf{b} = \{b_i\}$ be a binary sequence generated by the following recurrence relation

$$b_{n+k} = 1 \oplus b_k \oplus f_d(x_k),\ x_k = (b_{r_1+k}, \ldots, b_{r_t+k}) \in \mathbb{F}_{2^t}, d \in D_t^*,$$
$$0 < t < n,\ k \geq 0. \tag{5}$$

Similarly, for a proper selection of a t-tap position and a feedback function $f_d(x)$, the complementary binary sequence $\bar{\mathbf{b}}$ of $\mathbf{b}$ can be a *span n sequence*, but the sequence $\mathbf{b}$ is not a span n sequence since it contains the all-zero state.

If the number of terms in the algebraic normal form representation of the function f_d is even, then the recurrence relations (4) and (5) cannot generate a span n sequence for any choice of a t-tap position, since for the all-one state, recurrence relation (4) generates the all-one sequence, and recurrence relation (5) contains the all-one n-tuple.

Proposition 1 *If $f_d(x) = 0$ for $x = (1, 1, \ldots, 1) \in \mathbb{F}_{2^t}$, then recurrence relations (4) and (5) cannot generate span n sequences.*

In the recurrence relations (4) and (5), by varying three parameters, namely, the primitive polynomial $p(x)$, the decimation number d, and the t-tap position $(r_1, r_2, \ldots, r_t)$, a number of new span n sequences can be produced, and that number mainly depends on the length n of the NLFSR and the number t of inputs to the function f_d. We call this searching technique a *structured search*, where an NLFSR has a compact representation in terms of feedback functions and tap positions. Note that we may not always obtain a span n sequence for a fixed value of t and for any length n of the NLFSR. A special case of the recurrence relation (4) with the trace function in $(n-1)$ variables as the feedback function is defined in [37].

A periodic reverse binary sequence is defined as follows [32, 35]: for a binary sequence $\{a_0, a_1, \ldots, a_{2^n-2}\}$ with period $2^n - 1$, the reverse sequence of the binary sequence is defined by $\{a_{2^n-2}, a_{2^n-3}, \ldots, a_1, a_0\}$. A reverse sequence of a span n sequence is also a span n sequence, which is not shift equivalent to the original one, and the reverse span n sequence can be generated by the same function but with a different t-tap position.

Proposition 2 ([32]) *Let $g(x_0, x_1, \ldots, x_{n-1}) = x_0 \oplus f(x_1, \ldots, x_{n-1})$ generates a span n sequence with period $2^n - 1$. Then the function $h(x_0, x_{n-1}, \ldots, x_1) = x_0 \oplus f(x_{n-1}, \ldots, x_1)$ generates a reverse span n sequence.*

Our span n sequences generated by recurrence relations (4) and (5) with a fixed $P(x)$ are uniquely determined by the following three parameters:

1. the decimation number d,
2. the primitive polynomial $p(x)$,
3. the t-tap position $(r_1, r_2, \ldots, r_t)$.

Similarly, the reverse span n sequence of a span n sequence with parameters d, $p(x)$, and $(r_1, r_2, \ldots, r_t)$ is represented by the same decimation number d and the same primitive polynomial $p(x)$, but with a different t-tap position $(n - r_1, n - r_2, \ldots, n - r_t)$. For a fixed function $f_d(x)$, a span n sequence generated by $f_d(x)$ is different if the t-tap position is different. We now describe the span n sequence generation by the above structured search in the following example.

Example 1 The following example describes our span n sequence generation procedure for $t = 5$.

The WG transformation over $\mathbb{F}_{2^5}$ is given by

$$f(x) = \mathrm{Tr}(x + (x+1)^5 + (x+1)^{13} + (x+1)^{19} + (x+1)^{21}).$$

After simplification, $f(x)$ can be written as

$$f(x) = \mathrm{Tr}(x^{19}),\ x \in \mathbb{F}_{2^5},$$

which is degenerated into an m-sequence. For $t = 5$, the set of coset leaders is given by $D_t = \{1, 3, 5, 7, 11, 15\}$, and the coset leaders for which $f_d(x)$ is nonlinear is given by $D_t^* = \{1, 3, 7, 11, 15\}$, since for $d = 5$, the function $f_d(x)$ is linear. The d-th decimation of $f(x)$ is given by

$$f_d(x) = f(x^d) = \mathrm{Tr}(x^{d'}),\ d' = (19 \cdot d) \bmod 2^t - 1,\ d \in D_t^*.$$

The n-stage nonlinear recurrence relation with a t-tap position is given by

$$a_{n+k} = a_k \oplus f_d(x_k), x_k = (a_{r_1+k}, \ldots, a_{r_5+k}) \in \mathbb{F}_{2^5},\ k \geq 0.$$

The Boolean representation of $f(x) = \mathrm{Tr}(x^{19})$ with defining polynomial $p(x) = 1 + x + x^2 + x^4 + x^5$ of $\mathbb{F}_{2^5}$ is as follows:

$$\begin{aligned} f(x_0, \ldots, x_4) = {} & x_0 + x_3 + x_0x_1 + x_0x_2 + x_0x_3 + x_0x_4 + x_1x_2 + x_1x_3 + x_1x_4 \\ & + x_2x_4 + x_0x_1x_3 + x_0x_1x_4 + x_0x_2x_3 + x_0x_3x_4 + x_1x_2x_4. \end{aligned}$$

For the span n sequence with parameters $d = 1$, $p(x) = 1 + x + x^2 + x^4 + x^5$, $(r_1, r_2, r_3, r_4, r_5) = (1, 2, 3, 4, 5)$ in Table 4, the above recurrence relation can be written as

$$\begin{aligned} a_{7+k} = {} & a_k + a_{1+k} + a_{4+k} + a_{1+k}a_{2+k} + a_{1+k}a_{3+k} + a_{1+k}a_{4+k} + a_{1+k}a_{5+k} \\ & + a_{2+k}a_{3+k} + a_{2+k}a_{4+a} + a_{2+}a_{5+k} + a_{3+k}a_{5+k} + a_{1+k}a_{2+k}a_{4+k} \\ & + a_{1+k}a_{2+k}a_{5+k} + a_{1+k}a_{3+k}a_{4+k} + a_{1+k}a_{4+k}a_{5+k} + a_{2+k}a_{3+k}a_{5+k}, \\ & a_k \in \mathbb{F}_2, k \geq 0. \end{aligned}$$

Table 4 Span n sequences generated using WG5 for $n = 7$

Decimation	Polynomial	t-tap position
By recurrence relation (4)		
d	$(c_0, c_1, c_2, c_3, c_4)$	$(r_1, r_2, r_3, r_4, r_5)$
1	1 1 1 0 1	1 2 3 4 5
1	1 1 0 1 1	1 3 4 5 6
7	1 0 0 1 0	1 2 3 4 6
7	1 0 1 0 0	1 2 4 5 6
7	1 0 1 1 1	2 3 4 5 6
11	1 0 0 1 0	1 2 4 5 6
11	1 1 1 1 0	1 2 4 5 6
11	1 1 1 0 1	1 2 4 5 6
15	1 1 1 1 0	1 2 4 5 6
By recurrence relation (5)		
1	1 1 1 1 0	1 2 3 4 5
1	1 1 1 0 1	1 3 4 5 6
1	1 0 1 0 0	1 3 4 5 6
7	1 0 1 1 1	1 2 3 4 5
7	1 0 1 0 0	1 2 3 4 5
7	1 1 0 1 1	1 2 3 5 6
15	1 1 1 1 0	1 2 3 4 5

The above generates the following span n sequence of period $2^7 - 1$

$$11111110001110010001000001101100000010010110111010111000010111 1$$

$$01101010110010100001111001100010101001001111101001101000110011 10.$$

For $n = 7$, all the span n sequences produced by recurrence relations (4) and (5) are presented in Table 4.

4.2 Approximate Number of Functions in the Search Space

Note that three parameters, namely, a decimation number d, a primitive polynomial $p(x)$, and a t-tap position, determine a nonlinear recurrence relation or a feedback function that may generate a span n sequence. In other words, each feedback function can be considered as a candidate span n sequence. For a fixed value of n and t, a search space is formed by including all possible combinations of these three parameters. In order to find span n sequences, an exhaustive search is performed over this search space. We determine the size of the search space or the number of candidate span n sequences in terms of n and t in the following proposition.

Proposition 3 *For any $n > t \geq 6$, the number of feedback functions in the search space of recurrence relations (4) and (5) is given by $C = \left(\frac{\phi(2^t-1)}{t}\right)^2 \binom{n-1}{t}$ if $|D_t^*| = \frac{\phi(2^t-1)}{t}$.*

Proof As in the recurrence relations, the first position is fixed for the sequence to be periodic, and any t-tap position is chosen from $n-1$ positions ($n \geq 6$) to form a t-tap position; the number of distinct t-tap positions is given by $T = \binom{n-1}{t}$. Again, the total number of nonlinear feedback functions is given by $n_p \cdot |D_t^*|$, where $n_p = \frac{\phi(2^t-1)}{t}$ is the number of t degree primitive polynomials over $\mathbb{F}_2$ and $|D_t^*|$ is the number of decimation numbers for which the feedback function is nonlinear. Hence, for fixed n and t, the number of feedback functions in the search space is

$$C = n_p \cdot |D_t^*| \cdot T = \left(\frac{\phi(2^t-1)}{t}\right)^2 \binom{n-1}{t} \text{ if } |D_t^*| = \frac{\phi(2^t-1)}{t}.$$

□

Proposition 4 *A feedback shift register defined by recurrence relations (4) and (5) produces the maximum number of span n sequences when about half the length of the shift register tap positions participate in the feedback functions.*

Proof Without loss generality, we assume that the number of terms in a feedback function is even. In a feedback shift register, the feedback functions are different for different t-tap positions. Thus, for a particular value of n and t and for a feedback function in t variables, the number of different feedback functions in n variables is equal to $N_{n,t} = \binom{n-1}{t}$ and $N_{n,t}$ is maximum when $t = \lceil \frac{n}{2} \rceil$ (for linear feedback functions, t is always odd and $t \approx \lceil \frac{n}{2} \rceil$). If the feedback functions in n variables that are candidate span n sequences are uniformly distributed over the set of all Boolean functions, then the FSR generates the maximum number of span n sequences when $t \approx \lceil \frac{n}{2} \rceil$. Hence, the assertion is established. □

We note that an LFSR also produces the maximum number of span n sequences when $t \approx \lceil \frac{n}{2} \rceil$ (see Table 20). This property is also satisfied by the nonlinearly generated span n sequences using recurrence relations (4) and (5) (see Tables 6, 7, 8, 9, 10, 11, and 12). We now estimate the number of feedback functions in the search space for finding the maximum number of span n sequences. Assume that we use NLFSRs defined by recurrence relations (4) and (5) for $t = \lceil \frac{n}{2} \rceil$. Let N denote the number of span n sequences (including reverse span n sequences) obtained by recurrence relations (4) and (5). Then we have the following theorem.

Theorem 1 *An approximate number of candidate span n sequences or feedback functions in recurrence relations (4) and (5) is given by C_0, where $C_0 \approx \left(\frac{\phi(2^{\lceil \frac{n}{2} \rceil}-1)}{\lceil \frac{n}{2} \rceil}\right)^2 \cdot \frac{2^{n-1}}{\sqrt{\pi \cdot \frac{n-1}{2}}}$ and $C_0 \approx \frac{2^{2n-1}-2^{\frac{3n}{2}+1}}{\sqrt{\pi} \cdot (\lceil \frac{n}{2} \rceil)^{5/2}}$, if $2^t - 1$ is a Mersenne prime, and the success probability of obtaining such a span n sequence is given by $\frac{N}{C_0}$.*

Proof We recall that the size of the search space is

$$C = \left(\frac{\phi(2^t-1)}{t}\right)^2 \binom{n-1}{t}, \quad \text{for } |D_t^*| = \frac{\phi(2^t-1)}{t}.$$

Putting $t = \lceil \frac{n}{2} \rceil$ in the above formula, then we get

$$\begin{aligned} C_0 &= \left(\frac{\phi(2^{\lceil \frac{n}{2} \rceil}-1)}{\lceil \frac{n}{2} \rceil}\right)^2 \cdot \binom{n-1}{\lceil \frac{n}{2} \rceil} \\ &= \left(\frac{\phi(2^{\lceil \frac{n}{2} \rceil}-1)}{\lceil \frac{n}{2} \rceil}\right)^2 \cdot \binom{n-1}{\lfloor \frac{n-1}{2} \rfloor + 1}, \text{ for positive } n \\ &= \left(\frac{\phi(2^{\lceil \frac{n}{2} \rceil}-1)}{\lceil \frac{n}{2} \rceil}\right)^2 \cdot \frac{(n - \lfloor \frac{n-1}{2} \rfloor - 1) \cdot \binom{n-1}{\lfloor \frac{n-1}{2} \rfloor}}{(\lfloor \frac{n-1}{2} \rfloor + 1)}. \end{aligned}$$

By Stirling's formula

$$\binom{m}{\lfloor \frac{m}{2} \rfloor} \sim \frac{2^m}{\sqrt{\pi m/2}},$$

the above equation can be written as

$$\begin{aligned} C_0 &\sim \left(\frac{\phi(2^{\lceil \frac{n}{2} \rceil}-1)}{\lceil \frac{n}{2} \rceil}\right)^2 \cdot \frac{\lfloor \frac{n-1}{2} \rfloor \cdot 2^{n-1}}{(\lfloor \frac{n-1}{2} \rfloor + 1) \cdot \sqrt{\pi \cdot \frac{n-1}{2}}} \\ &\sim \left(\frac{\phi(2^{\lceil \frac{n}{2} \rceil}-1)}{\lceil \frac{n}{2} \rceil}\right)^2 \cdot \frac{2^{n-1}}{\sqrt{\pi \cdot \frac{n-1}{2}}}. \\ &\approx \frac{2^{2n-1} - 2^{\frac{3n}{2}+1}}{\sqrt{\pi} \cdot (\lceil \frac{n}{2} \rceil)^{5/2}}, \text{ if } 2^t - 1 \text{ is a Mersenne prime.} \end{aligned}$$

Thus, the success probability of obtaining a span n sequence is equal to $\frac{N}{C_0}$. Hence, the result is proved. □

5 Experimental Results on Span n Sequence Generation Using WG Transformations

In this section, we report the number of new span n sequences generated using WG transformations. We also present a heuristic method for searching WG span n sequences of long length. Table 5 provides a summary of the list of orthogonal functions used to produce span n sequences.

5.1 WG Span n Sequences

WG span n sequences are obtained by putting the WG transformation in recurrence relations (4) and (5) for different t and n. The span n sequences are generated by computer simulations. We consider the WG transformations over the field $\mathbb{F}_{2^t}$ for $t = 5, 7, 8, 10$, and 11. We denote by WG-t the WG transformations over the field $\mathbb{F}_{2^t}$. Table 6 presents the number of new span n sequences (new reverse span n sequences are not taken into account) produced by recurrence relations (4) and (5) for $6 \leq n \leq 20$. However, this method can be applied to generate span n sequences of long length. In Table 6, "$\times$" denotes the recurrence relations that are not defined for such values of n and t, and $\sim$ represents those cases wherein the number of span n sequences is not yet determined. We present some instances of new span n sequences in the Appendix and all span n sequences in http://www.comsec.uwaterloo.ca/~kmandal/WG-Span-n/index.html.

A graphical representation of the number of new span n sequences is provided in Fig. 1, which shows that for different t the distribution of the number of new span n sequences has the following property: the number of span n sequences increases as n increases, and it reaches the maximum for some value of n, and thereafter the number of span n sequences decreases as n increases. At a quick glance, we can observe that the number of span n sequences is maximal close to $n = 2t$, which follows from the fact that the size of the search space is a multiple of a binomial coefficient (see Proposition 4). This fact reveals that there exists a trade off between n and t for obtaining the maximum number of span n sequences.

Table 5 Orthogonal functions used in the structured search

Parameter t	Orthogonal functions
$t = 5$	T1, T3
$t = 7$	T1, T3, T5, WG, Hall, QR
$t = 8$	T5, WG, GMW
$t = 9$	T1, T3, GMW, Segre, Glynn 1
$t = 10$	T1, T5, WG, GMWi, $i = 1 \ldots 7$
$t = 11$	T1, T3, T5, WG, Segre, Glynn 1, Glynn 2, B3

Table 6 Number of WG span n sequences

By recurrence relation (4)															
	n														
t	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	0	9	7	14	8	11	17	11	13	10	3	7	7	0	1
7	×	×	3	25	42	63	108	138	138	125	126	111	83	86	63
8	×	×	×	3	9	18	34	76	96	104	106	108	110	90	79
10	×	×	×	×	×	5	40	107	246	373	627	819	999	∼	∼
11	×	×	×	×	×	×	31	204	574	1313	2539	4079	∼	∼	∼
Total	0	9	10	42	59	97	230	536	1067	1925	3401	5124	–	–	–
By recurrence relation (5)															
t	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	1	7	7	10	16	18	10	8	4	10	2	1	3	1	0
7	×	×	4	25	47	59	121	122	137	125	123	98	74	84	54
8	×	×	×	1	6	35	33	75	73	91	123	115	106	99	77
10	×	×	×	×	×	4	47	118	270	401	680	863	∼	∼	∼
11	×	×	×	×	×	×	33	186	576	1350	2522	4010	∼	∼	∼
Total	1	7	11	36	69	116	244	509	1060	1977	3450	5087	–	–	–

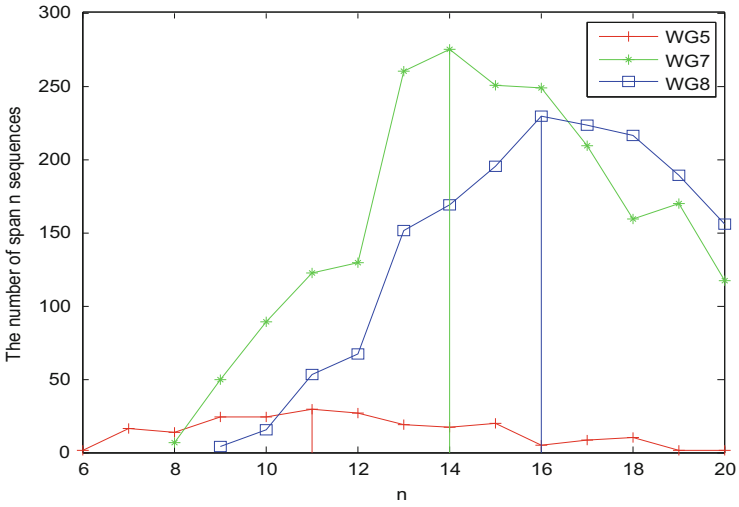

Fig. 1 Distribution of the number of span n sequences

Remark 1 There exist many span n sequences whose t-tap positions and the bases of the finite fields are the same, but their decimation numbers are different.

5.2 The Search Complexity Reduction for WG Span n Sequences

It is worth noticing that as t increases, the number of feedback functions in the search space increases exponentially. For large t, it is hard to find span n sequences by considering all functions in the search space. Thus, for large n and t, a search in a restricted search space can be performed to find span n sequences by imposing restrictions over decimation numbers and t-tap positions. Below we list a type of decimation numbers and t-tap positions that are observed for WG span n sequences. In some cases, we may not find any span n sequence. However, according to our observations, it is possible to obtain many span n sequences.

5.2.1 Observations on Decimation Numbers

We have performed a search on the following type of decimation numbers for different n

$$D_{\text{dec}} = \{d : d \in D_t^* \text{ and } d = 2^i - 1, i = 1, 2, \ldots, t-1\}$$

for $t = 7$, 8, and 10, and the result shows that there exist many span n sequences whose decimation numbers in the recurrence relations (4) and (5) are of the above type. For this type of decimation numbers in the recurrence relations, the size of the search space is given by

$$C_{\text{dec}} = \frac{\phi(2^t - 1)}{t}(t-1)\binom{n-1}{t} \approx \phi(2^t - 1)\binom{n-1}{t}.$$

Obviously, the reduced complexity C_{dec} is less than the original complexity C.

5.2.2 Observations on t-Tap Positions

Likewise, a search in the search space can be performed according to some pattern of t-tap positions for finding long period span n sequences. Assume that it is possible to fix, say, k tap positions ($1 \le k \le t$). Then, the total number of fixed tap positions in the recurrence relations is $(k+1)$, and we only need to choose $(t-k)$ positions out of $(n-1-k)$ positions. So, for k fixed choices of tap positions, the search complexity is

$$C_{\text{tap}} = \left(\frac{\phi(2^t - 1)}{t}\right)^2 \binom{n-1-k}{t-k}.$$

Based on our observations on the t-tap positions for $t = 7, 8,$ and 10, the following types of t-tap positions are effective when the slope of the curves in Fig. 1 increases gradually. For example, when $t = 7$, $n = 11, 12, 13,$ and 14 and $t = 8$, $n = 13, 14, 15, 16, 17,$ and 18, the t-tap positions are given by: $\{1, 2, 3, 4, \ldots\}$, $\{1, 2, 3, \ldots, n-1\}$, $\{1, 2, \ldots, n-2, n-1\}$, $\{1, \ldots, n-3, n-2, n-1\}$, where the numbers in the tap positions represent fixed positions in the t-tap positions (i.e., $k = 4$ fixed positions) and "..." represents a combination of $(n-k-1)$ tap positions. We performed a search according to the first pattern of t-tap position; the following span n sequence generated by a WG transformation has been found for $t = 13$ and $n = 24$.

Decimation d	Polynomial $(c_0, c_1, c_2, \ldots, c_{11}, c_{12})$	t-tap position $(r_1, r_2, \ldots, r_{12}, r_{13})$
1207	(1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0)	(1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 22)

6 Experimental Results on Span n Sequences Generated by Other Orthogonal Functions

This section reports the number of span n sequences produced using three-term, five-term, monomial, Hall, quadratic residue, Glynn, Segre, GMW, and Kasami power functions. Explicit representations of these function are provided in Tables 1, 2, and 3.

6.1 Three-Term and Five-Term and Monomial Span n Sequences

Considering three-term and five-term functions in recurrence relations (4) and (5), a number of span n sequences can be obtained by the structured search. Tables 7 and 8 present the number of span n sequences for three-term functions and five-term functions, respectively. When $t = 5$, three-term functions and five-term functions degenerate to the same functions, as a result, the number of span n sequences obtained by three-term functions and five-term function are the same.

Table 9 presents the number of span n sequences produced using monomial functions for $6 \leq n \leq 20$. In tables, $\times$ denotes that the recurrence relation is not defined by the parameters t and n, and $\sim$ denotes that the cases are incomplete due to a huge number of functions in the search space. When $t = 5$, the WG transformations and monomial functions degenerate to the same functions.

Table 7 Number of three-term span n sequences

By recurrence relation (4)												
	n											
t	6	7	8	9	10	11	12	13	14	15	16	17
5	1	3	9	8	9	8	4	3	5	2	3	1
7	×	×	6	25	51	89	103	150	131	128	127	123
9	×	×	×	×	8	52	104	223	391	549	710	770
11	×	×	×	×	×	×	35	190	624	1323	2580	4056
Total	1	3	15	33	68	149	246	566	1151	2002	3420	4950
By recurrence relation (5)												
t	6	7	8	9	10	11	12	13	14	15	16	17
5	1	2	2	5	10	5	6	5	3	1	3	5
7	×	×	4	24	44	84	98	122	133	146	128	111
9	×	×	×	×	12	47	109	237	361	553	694	823
11	×	×	×	×	×	×	34	186	578	1416	2554	4007
Total	1	3	6	29	66	136	247	550	1075	2116	3379	4946

Table 8 Number of five-term span n sequences

By recurrence relation (4)														
	n													
t	6	7	8	9	10	11	12	13	14	15	16	17	18	19
5	1	3	9	8	9	8	4	4	5	2	3	1	0	1
7	×	×	5	22	44	66	118	131	115	135	124	118	99	90
8	×	×	×	1	9	18	37	56	88	101	104	86	92	90
10	×	×	×	×	×	9	37	116	246	411	621	797	943	∼
11	×	×	×	×	×	×	25	171	590	1443	2618	4194	∼	∼
Total	1	3	14	31	62	101	221	478	1044	2092	3470	5196	–	–
By recurrence relation (5)														
t	6	7	8	9	10	11	12	13	14	15	16	17	18	19
5	1	2	2	5	10	5	6	5	3	1	3	5	0	1
7	×	×	8	19	43	74	108	138	138	127	117	102	84	91
8	×	×	×	0	6	22	38	54	66	116	89	106	83	93
10	×	×	×	×	×	7	47	119	223	443	627	861	∼	∼
11	×	×	×	×	×	×	20	172	609	1397	2558	4062	∼	∼
Total	1	2	10	24	59	108	219	488	1039	2084	3394	5136	–	–

6.2 Hall, QR, Segre, Glynn, and GMW Span n Sequences

In this section, we present the number of span n sequences produced by Hall, QR, Segre, Glynn, and GMW functions for $7 \leq n \leq 20$. We use the functions defined in Tables 1, 2, and 3 for Hall, quadratic residue, Glynn, Segre, and GMW functions in recurrence relations (4) and (5).

Table 9 Number of span n sequences generated by monomial functions

By recurrence relation (4)														
	n													
t	6	7	8	9	10	11	12	13	14	15	16	17	18	19
5	0	9	7	14	8	11	17	11	13	10	3	7	7	0
7	×	×	6	17	41	76	79	118	108	99	125	78	88	72
9	×	×	×	×	10	43	120	258	410	519	662	788	∼	∼
11	×	×	×	×	×	×	26	188	604	1423	2491	4056	∼	∼
Total	0	9	13	31	59	130	242	575	1135	2051	3281	4929	–	–
By recurrence relation (5)														
t	6	7	8	9	10	11	12	13	14	15	16	17	18	19
5	1	7	7	10	16	18	10	8	4	10	2	1	3	1
7	×	×	4	25	45	60	98	117	114	104	116	96	86	77
9	×	×	×	×	6	37	131	239	367	558	740	860	∼	∼
11	×	×	×	×	×	×	32	184	596	1403	2547	4074	∼	∼
Total	1	7	11	35	67	115	271	548	1081	2075	3405	5031	–	–

Table 10 Number of span n sequences generated by Hall functions and QR functions

By recurrence relation (4)														
		n												
t	OF	8	9	10	11	12	13	14	15	16	17	18	19	20
7	Hall	2	9	19	21	41	38	35	45	28	34	30	30	–
7	QR	0	4	4	5	14	27	16	9	18	14	12	6	6
By recurrence relation (5)														
t	OF	8	9	10	11	12	13	14	15	16	17	18	19	20
7	Hall	1	6	20	25	37	48	36	44	46	24	39	–	–
7	QR	0	3	6	7	13	12	13	18	16	13	14	10	8

For the range $7 \leq t \leq 11$, the Hall and QR functions with trace representations exist only for $t = 7$. Table 10 presents the number of span n sequences produced using recurrence relations (4) and (5) with Hall and QR functions for $8 \leq n \leq 20$. When all the decimated QR functions are considered, the class of 18 QR functions degenerates to two distinct QR orthogonal functions, and similarly, the class of 18 Hall functions degenerates to six distinct Hall orthogonal functions. Due to this reason, the number of span n sequences in Table 10 is smaller compared to other cases for $n = 7$.

When all the decimations are considered, Glynn 1 functions and Glynn 2 functions over $\mathbb{F}_{2^9}$ degenerate to the same class of orthogonal functions. Therefore, the number of span n sequences for Glynn 1 and Glynn 2 functions are the same in the structured search. However, for $t = 11$, the Glynn 1 class of functions and Glynn 2 class of functions are different. We provide the number of span n sequences produced by Glynn functions in Table 11, which also contains the number of span n sequences generated by Segre functions for $t = 9$ and 11. In Tables 10, 11, and 12,

Table 11 Number of span n sequences generated by Segre and Glynn functions

By recurrence relation (4)										
		n								
t	OF	10	11	12	13	14	15	16	17	18
9	Segre	15	51	131	245	418	528	706	783	–
11	Segre	×	×	34	172	586	1413	2564	–	–
9	Glynn 1	11	52	129	253	415	584	673	790	–
11	Glynn 1	×	×	28	177	587	1418	2553	–	–
11	Glynn 2	×	×	30	185	595	1320	2646	–	–
By recurrence relation (5)										
t	OF-t	10	11	12	13	14	15	16	17	18
9	Segre	7	48	108	264	371	521	692	–	–
11	Segre	×	×	37	153	627	1372	–	–	–
9	Glynn 1	6	49	126	248	397	529	709	–	–
11	Glynn 1	×	×	26	185	562	1351	–	–	–
11	Glynn 2	×	×	28	183	598	1340	–	–	–

Table 12 Number of span n sequences generated by GMW functions

By recurrence relation (4)												
		n										
t	OF	9	10	11	12	13	14	15	16	17	18	# of terms
8	GMW	1	11	13	50	75	71	99	97	117	78	4
9	GMW	×	15	45	128	223	382	–	–	–	–	3
10	GMW1	×	×	7	37	114	236	424	606	810	–	2
10	GMW2	×	×	6	51	97	247	405	–	–	–	2
10	GMW3	×	×	5	33	119	255	415	672	865	–	4
10	GMW4	×	×	7	36	110	248	405	–	–	–	4
10	GMW5	×	×	10	39	147	261	411	645	853	–	7
10	GMW6	×	×	5	39	113	234	440	654	816	–	8
10	GMW7	×	×	10	39	118	236	422	664	888	–	14
By recurrence relation (5)												
t	OF	9	10	11	12	13	14	15	16	17	18	# of terms
8	GMW	1	5	21	45	77	80	90	107	116	111	4
9	GMW	×	11	44	140	247	414	559	716	–	–	3
10	GMW1	×	×	7	34	117	257	414	609	–	–	2
10	GMW2	×	×	8	41	126	243	409	–	–	–	2
10	GMW3	×	×	7	44	122	257	411	641	–	–	4
10	GMW4	×	×	4	35	130	257	424	–	–	–	4
10	GMW5	×	×	6	43	113	239	407	638	–	–	7
10	GMW6	×	×	2	42	113	247	455	630	–	–	8
10	GMW7	×	×	5	51	133	258	429	643	–	–	14

"–" denotes the computation for the number of span n sequences is in progress and will be finished soon.

Table 12 presents the number of span n sequences produced by GMW functions in the structured search for $t = 8, 9$, and 10 and $9 \leq n \leq 19$. For the GMW functions over $\mathbb{F}_{2^8}$ and $\mathbb{F}_{2^9}$, there exists only one class of GMW functions. On the other hand, for the GMW functions over $\mathbb{F}_{2^{10}}$, there exist total seven distinct classes of orthogonal GMW functions with different number of terms in the trace representation. GMW span n sequences with $9 \leq n \leq 18$ are generated using recurrence relations (4) and (5) with GMWi functions, $1 \leq i \leq 7$. In Table 12, the term "# of terms" denotes the number of terms in the trace representation of a GMW function.

Remark 2 For a class of orthogonal functions in recurrence relations (4) and (5), each span n sequence is uniquely determined by a decimation number, a primitive polynomial, and a t-tap position. Unfortunately, we could not find any relation among these three parameters.

7 The Success Probability Comparison

In this section, an empirical success probability of obtaining a span n sequence using a orthogonal feedback function is presented. Note that the success probability of obtaining a randomly generated span n sequence is $\frac{1}{2^{n-3}}$ [33], where a random span n sequence is generated by randomly choosing a feedback function from the set of all Boolean functions in n variables and checking the condition for a span n sequence.

We compared the success probability of obtaining a span n sequence using WG transformations (including reverse sequences) in the structured search with a random span n sequence generation method for $t = 5, 7, 8$ (for $t \approx \lceil \frac{n}{2} \rceil$), 10, and 11 (for $13 \leq n \leq 17$), and the comparison shows that in the structured search, one can produce a span n sequence with a better success probability than that of a random span n sequence generation method. A comparison of success probability for $t = 5, 7,$ and 8 is provided in Table 13. Furthermore, we compared the success probability of obtaining a span n sequences using three-term, five-term, and monomial functions in Table 13 for $t = 5, 7, 8, 9$. Table 13 illustrates that a span n sequence can be produced using any of three-term, five-term, and monomial functions with a better success probability. Our empirical comparisons also show that the success probability of obtaining a span n sequence using Hall, QR, Segre, Glynn, and GMW functions is greater than that of a random span n sequence generation method. We don't provide the success probability values due to the large number of cases.

Table 13 The success probability comparison for WG, three-term, five-term, and monomial span n sequences

WG span n sequences			
	$n = 2t$	Our approach	Randomly chosen
WG-5	10	$\frac{1}{2^{6.56}}$	$\frac{1}{2^7}$
WG-7	14	$\frac{1}{2^{9.98}}$	$\frac{1}{2^{11}}$
WG-8	16	$\frac{1}{2^{11.81}}$	$\frac{1}{2^{13}}$
Three-term span n sequences			
	$n \approx 2t$	Our approach	Randomly chosen
T3-5	10	$\frac{1}{2^{6.89}}$	$\frac{1}{2^7}$
T3-7	14	$\frac{1}{2^{10.04}}$	$\frac{1}{2^{11}}$
T3-9	17	$\frac{1}{2^{13.04}}$	$\frac{1}{2^{14}}$
Five-term span n sequences			
	$n = 2t$	Our approach	Randomly chosen
T5-5	10	$\frac{1}{2^{6.89}}$	$\frac{1}{2^7}$
T5-7	14	$\frac{1}{2^{10.10}}$	$\frac{1}{2^{11}}$
T5-8	16	$\frac{1}{2^{12.02}}$	$\frac{1}{2^{13}}$
Monomial span n sequences			
	$n \approx 2t$	Our approach	Randomly chosen
T1-5	10	$\frac{1}{2^{6.88}}$	$\frac{1}{2^7}$
T1-7	14	$\frac{1}{2^{10.29}}$	$\frac{1}{2^{11}}$
T1-9	17	$\frac{1}{2^{12.96}}$	$\frac{1}{2^{14}}$

8 Linear Span of New Span n Sequences

In this section, we analyze the linear span of new span n sequences produced by orthogonal functions and present two conjectures on linear span of span n sequences produced by orthogonal functions.

We study the linear span of new span n sequences generated using orthogonal functions. The linear span of a sequence is an important randomness property that is considered as an upper bound on sequence unpredictability because using only twice-linear span consecutive bits one can certainly predict the remaining bits of the sequence by the Berlekamp–Massey algorithm [2, 31]. Sequences with optimal linear complexity are of practical interests, since an attacker requires the whole sequence to decrypt the message in a stream cipher. There is no theoretical result on the linear span of span n sequences generated by a nonlinear feedback shift register. What we know is the bounds presented in Property 1 in Sect. 2.

We compute the linear span of new span n sequences by the Berlekamp–Massey algorithm, and our computational results show that the linear span of a new sequence lies in the range of $(2^n - 2 - 3n)$ (near optimal) and $(2^n - 2)$ (optimal). Table 14 presents a summary of the linear spans of WG span n sequences generated by the recurrence relations (4) and (5), respectively. Moreover, Tables 15, 16, and 17 exhibit a summary of the linear spans of the span n sequences generated by

Table 14 The bounds of the linear span of WG span n sequences

Range on n	t	Upper bound of LS	Lower bound of LS
By recurrence relation (4)			
$7 \le n \le 20$	5	$2^n - 2$	$2^n - 2 - 2n$
$8 \le n \le 20$	7	$2^n - 2$	$2^n - 2 - 2n$
$9 \le n \le 20$	8	$2^n - 2$	$2^n - 2 - 3n$
$11 \le n \le 17$	10	$2^n - 2$	$2^n - 2 - 3n$
$12 \le n \le 17$	11	$2^n - 2$	$2^n - 2 - 2n$
By recurrence relation (5)			
$7 \le n \le 20$	5	$2^n - 2$	$2^n - 2 - 2n$
$8 \le n \le 20$	7	$2^n - 2$	$2^n - 2 - 3n$
$9 \le n \le 20$	8	$2^n - 2$	$2^n - 2 - 3n$
$11 \le n \le 17$	10	$2^n - 2$	$2^n - 2 - 3n$
$12 \le n \le 16$	11	$2^n - 2$	$2^n - 2 - 3n$

Table 15 The bounds of the linear span of monomial span n sequences

Range on n	t	Upper bound of LS	Lower bound of LS
By recurrence relation (4)			
$7 \le n \le 19$	5	$2^n - 2$	$2^n - 2 - 2n$
$8 \le n \le 19$	7	$2^n - 2$	$2^n - 2 - 3n$
$8 \le n \le 17$	9	$2^n - 2$	$2^n - 2 - 3n$
$12 \le n \le 16$	11	$2^n - 2$	$2^n - 2 - 3n$
By recurrence relation (5)			
$7 \le n \le 19$	5	$2^n - 2$	$2^n - 2 - 2n$
$8 \le n \le 19$	7	$2^n - 2$	$2^n - 2 - 3n$
$8 \le n \le 17$	9	$2^n - 2$	$2^n - 2 - 3n$
$12 \le n \le 16$	11	$2^n - 2$	$2^n - 2 - 3n$

Table 16 The bounds of the linear span of three-term span n sequences

Range on n	t	Upper bound of LS	Lower bound of LS
By recurrence relation (4)			
$7 \le n \le 17$	5	$2^n - 2$	$2^n - 2 - 2n$
$8 \le n \le 17$	7	$2^n - 2$	$2^n - 2 - 3n$
$8 \le n \le 17$	9	$2^n - 2$	$2^n - 2 - 3n$
$12 \le n \le 17$	11	$2^n - 2$	$2^n - 2 - 3n$
By recurrence relation (5)			
$7 \le n \le 17$	5	$2^n - 2$	$2^n - 2 - 2n$
$8 \le n \le 17$	7	$2^n - 2$	$2^n - 2 - 2n$
$8 \le n \le 17$	9	$2^n - 2$	$2^n - 2 - 3n$
$12 \le n \le 17$	11	$2^n - 2$	$2^n - 2 - 2n$

monomial functions, three-term functions, and five-term functions, respectively, for different values of t, and Table 18 presents a summary of the linear span of span n sequences produced by other orthogonal functions. Our computational results also show that most of new sequences obtain the optimal linear span $(2^n - 2)$, only very

Table 17 The bounds of the linear span of five-term span n sequences

Range on n	t	Upper bound of LS	Lower bound of LS
By recurrence relation (4)			
$7 \le n \le 19$	5	$2^n - 2$	$2^n - 2 - 2n$
$8 \le n \le 19$	7	$2^n - 2$	$2^n - 2 - 2n$
$9 \le n \le 19$	8	$2^n - 2$	$2^n - 2 - 3n$
$11 \le n \le 17$	10	$2^n - 2$	$2^n - 2 - 3n$
$12 \le n \le 16$	11	$2^n - 2$	$2^n - 2 - 2n$
By recurrence relation (5)			
$7 \le n \le 20$	5	$2^n - 2$	$2^n - 2 - 2n$
$8 \le n \le 20$	7	$2^n - 2$	$2^n - 2 - 3n$
$9 \le n \le 20$	8	$2^n - 2$	$2^n - 2 - 3n$
$11 \le n \le 17$	10	$2^n - 2$	$2^n - 2 - 2n$
$12 \le n \le 16$	11	$2^n - 2$	$2^n - 2 - 3n$

Table 18 The upper and lower bounds of the linear span of Hall, QR, GMW, Segre, and Glynn span n sequences

By recurrence relations (4) and (5)				
t	Function	Range on n	Upper bound	Lower bound
7	Hall	$8 \le n \le 19$	$2^n - 2$	$2^n - 2 - 2n$
	QR	$8 \le n \le 20$	$2^n - 2$	$2^n - 2 - 3n$
8	GMW	$9 \le n \le 18$	$2^n - 2$	$2^n - 2 - 2n$
9	Segre	$10 \le n \le 16$	$2^n - 2$	$2^n - 2 - 3n$
	Glynn	$10 \le n \le 16$	$2^n - 2$	$2^n - 2 - 3n$
	GMW	$10 \le n \le 16$	$2^n - 2$	$2^n - 2 - 3n$
10	GMW	$11 \le n \le 17$	$2^n - 2$	$2^n - 2 - 3n$
11	Segre	$12 \le n \le 16$	$2^n - 2$	$2^n - 2 - 3n$
	Glynn	$12 \le n \le 16$	$2^n - 2$	$> 2^n - 2 - 3n$

few span n sequences obtain the linear span $(2^n - 2 - 3n)$, and in some cases all the linear spans are greater than $(2^n - 2 - 3n)$.

Based on our observation on the linear span of new span n sequences produced by orthogonal functions, we have the following two conjectures. These two conjectures are valid and verified by our computational results for $n \le 20$.

Conjecture 1 Let the function g be an orthogonal function and $\mathbf{s} = \{s_i\}$ be a binary sequence generated by an n-stage NLFSR with $n > m$ whose feedback function is given by

$$f(x_0, x_1, \ldots, x_{n-1}) = c \oplus x_0 \oplus g(y)$$

where $c = 0/1$ and $y = (x_{r_1}, x_{r_2}, \ldots, x_{r_m}), y \in \mathbb{F}_{2^m}$, and $0 < r_1 < r_2 < \cdots < r_m < n$. If $\mathbf{s}$ or $\bar{\mathbf{s}}$ is a span n sequence, then the linear span of $\mathbf{s}$, denoted as LS_s, is bounded by

$$(2^n - 2 - 3n) \le LS_s \le (2^n - 2).$$

Conjecture 2 For a prime length of an NLFSR, the linear span of a span n sequence produced by the above feedback function with an orthogonal function takes one of the following three values $\{2^n - 2 - 2n, 2^n - 2 - n, 2^n - 2\}$.

9 Applications

Our span n sequences and span n sequences produced by the structured search in this chapter can be used in the following scenarios. In [28], Mandal and Gong analyzed the composited construction based on a span n sequence for generating long and strong de Bruijn sequences. Based on their analysis, the span n sequence to be used in the construction must have high linear span in order to produce strong de Bruijn sequences. Since our span sequences have optimal or near-optimal linear span, these span n sequences can be used in the composited construction for producing long and strong de Bruijn sequences. Mandal et al. [30] designed Warbler, a pseudorandom number generator for EPC C1 Gen2 RFID tags using NLFSRs where two span n sequences with optimal linear span are used to promise the randomness properties such as period and linear span of an output sequence. Our span n sequences or span n sequences produced by the structured search can be used to design lightweight pseudorandom number generators and stream ciphers. Thus, our span n sequences have an immediate application in cryptography, which can be found in [28, 30].

Conclusion
In this chapter, we have studied the span n sequence generation using orthogonal functions and presented some theoretical results on generating span n sequences and experimental results about the number of span n sequences produced by orthogonal functions. We used all known and well-studied orthogonal functions as nonlinear feedback functions in an NLFSR for $5 \leq t \leq 11$ and presented the number of span n sequences produced using orthogonal functions for $6 \leq n \leq 20$. Finally, we analyzed the linear span of new span n sequences produced by the orthogonal functions and gave a summary of the bounds of the linear span for each class of span n sequences. Interestingly, the linear span of a new span n sequence lies between the near optimal $(2^n - 2 - 3n)$ and optimal $(2^n - 2)$. We observed that the majority of span n sequences have an optimal linear span. According to our study, it is possible to obtain span n sequences of high linear span with a better probability of success using orthogonal feedback functions.

Appendix: A Upper and Lower Bounds of Linear Span of Span n Sequences

We present the upper and lower bounds of the linear span of new span n sequences generated using orthogonal functions for different n and t and give all new span n sequences generated using WG transformations for $t = 5$ (Tables 14, 15, 16, 17, 18, 19, 20, 21, and 22). All new span n sequences generated using WG transformations with $t = 7, 8, 10$, and 11 can be found in http://www.comsec.uwaterloo.ca/~kmandal/WG-Span-n/index.html.

Table 19 Span n sequences generated using WG7

Length n	Decimation d	Polynomial $(c_0, c_1, \ldots, c_5, c_6)$	t-tap position $(r_1, r_2, \ldots, r_6, r_7)$
8	5	1 1 0 0 0 0 0	1 2 3 4 5 6 7
9	1	1 0 1 1 1 1 1	1 2 3 4 5 6 7
10	27	1 1 1 1 0 1 1	1 2 3 4 5 6 7
11	1	1 1 1 1 0 1 1	1 2 3 5 8 9 10
12	1	1 0 1 1 1 0 0	1 2 4 5 8 10 11
13	9	1 1 0 0 1 0 1	1 2 3 4 5 6 8
14	43	1 1 1 0 1 1 1	1 2 3 4 5 6 7
15	31	1 1 0 0 0 0 0	1 2 3 4 7 12 14
16	27	1 1 1 1 0 1 1	1 2 3 5 6 8 14
17	1	1 0 1 1 1 0 0	1 2 3 4 7 9 13
18	1	1 0 1 1 1 0 0	1 2 3 4 6 9 16
19	3	1 1 1 1 1 1 0	1 2 3 5 7 15 17
20	31	1 1 1 1 1 1 0	1 2 3 7 8 12 15

Table 20 Tap-position distribution for an LFSR of length ≤ 20

# of taps	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	2	2	4	–	2	2	2	–	–	–	6	–	6	2	–	2
4	4	4	10	12	16	20	44	18	66	42	82	52	152	72	158	100
6	–	–	4	4	28	28	80	86	236	226	470	368	1050	718	1774	1104
8	–	–	–	–	2	10	50	36	264	338	720	812	2674	2296	6696	4522
10	–	–	–	–	–	–	–	4	60	140	450	648	2696	2910	10238	8436
12	–	–	–	–	–	–	–	–	4	12	66	156	1006	1470	6766	7000
14	–	–	–	–	–	–	–	–	–	–	6	12	122	284	1772	2460
16	–	–	–	–	–	–	–	–	–	–	–	–	–	24	190	354
18	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	22

Table 21 WG span n sequences generated using rec. rel. (4)

n	Decimation d	Polynomial $(c_0, c_1, c_2, c_3, c_4)$	Tap position $(r_1, r_2, r_3, r_4, r_5)$
8	1	1 0 1 0 0	1 2 4 5 7
	1	1 1 1 1 0	1 3 4 5 6
	1	1 1 1 1 0	2 4 5 6 7
	3	1 1 0 1 1	1 2 3 5 6
	7	1 0 1 1 1	1 2 3 5 7
	7	1 0 1 0 0	2 3 4 6 7
	15	1 1 1 1 0	2 3 4 6 7
9	1	1 1 1 0 1	1 2 5 6 8
	1	1 1 1 0 1	1 3 6 7 8
	1	1 1 1 1 0	2 3 5 7 8
	1	1 1 1 0 1	4 5 6 7 8
	3	1 1 0 1 1	1 2 4 5 6
	3	1 0 1 0 0	1 2 4 5 8
	3	1 0 1 0 0	2 4 6 7 8
	7	1 0 1 0 0	1 2 3 4 6
	11	1 1 1 0 1	1 4 6 7 8
	11	1 1 1 1 0	2 4 5 6 7
	11	1 1 1 1 0	2 4 5 6 8
	11	1 1 1 0 1	2 4 6 7 8
	15	1 1 1 1 0	1 2 3 4 6
	15	1 1 1 0 1	1 2 5 7 8
10	1	1 1 0 1 1	1 2 4 5 8
	1	1 1 1 0 1	1 3 4 6 7
	1	1 1 1 0 1	1 3 4 6 9
	3	1 1 0 1 1	1 2 3 4 8
	7	1 0 0 1 0	1 2 4 7 8
	11	1 0 1 1 1	1 2 3 4 5
	11	1 0 0 1 0	1 2 3 7 8
	11	1 1 1 1 0	1 4 5 8 9
11	1	1 1 1 0 1	1 2 7 8 10
	1	1 1 1 1 0	3 4 5 8 10
	1	1 1 1 0 1	6 7 8 9 10
	7	1 0 1 1 1	1 2 3 6 7
	7	1 0 0 1 0	1 3 7 8 10
	7	1 0 1 1 1	2 3 4 7 10
	7	1 1 0 1 1	2 3 7 9 10
	7	1 0 0 1 0	2 4 5 6 10
	7	1 1 0 1 1	3 4 5 8 9
	11	1 1 1 1 0	1 2 4 5 8
	11	1 1 1 0 1	1 3 4 6 10

(continued)

Table 21 (continued)

n	Decimation d	Polynomial $(c_0, c_1, c_2, c_3, c_4)$	Tap position $(r_1, r_2, r_3, r_4, r_5)$
12	1	1 1 1 1 0	2 3 4 5 6
	1	1 0 1 0 0	2 3 4 5 8
	1	1 1 1 0 1	2 3 5 7 9
	1	1 0 1 0 0	2 3 6 9 10
	1	1 1 1 0 1	4 6 9 10 11
	3	1 1 0 1 1	1 2 3 4 5
	3	1 1 0 1 1	2 5 7 8 10
	3	1 0 1 0 0	4 5 6 9 11
	7	1 0 1 0 0	1 2 4 7 8
	7	1 1 0 1 1	1 2 5 6 8
	11	1 0 0 1 0	1 3 4 6 10
	11	1 1 1 0 1	1 3 4 9 11
	11	1 1 1 1 0	1 4 5 8 9
	11	1 1 1 0 1	2 3 6 7 10
	11	1 1 1 1 0	3 5 7 8 9
	11	1 1 1 1 0	4 6 7 9 10
	15	1 1 1 1 0	1 2 4 7 8

Table 22 WG span n sequences generated using rec. rel. (4)

n	Decimation d	Polynomial $(c_0, c_1, c_2, c_3, c_4)$	Tap position $(r_1, r_2, r_3, r_4, r_5)$
13	1	1 0 1 0 0	1 3 4 5 9
	1	1 0 1 0 0	5 8 9 11 12
	3	1 1 0 1 1	5 6 10 11 12
	7	1 0 1 0 0	1 2 3 6 8
	7	1 1 0 1 1	3 5 7 10 12
	7	1 1 0 1 1	6 7 9 10 12
	11	1 0 0 1 0	1 2 3 5 10
	11	1 1 1 0 1	1 2 5 10 12
	11	1 1 1 0 1	1 5 6 10 12
	11	1 1 1 0 1	4 5 7 8 9
	15	1 1 1 1 0	1 2 3 6 8
14	1	1 0 1 0 0	1 3 5 7 9
	1	1 1 1 1 0	2 6 8 9 13
	1	1 1 1 0 1	3 4 6 8 10
	1	1 1 1 0 1	3 5 8 10 13
	3	1 1 0 1 1	1 8 10 11 13

(continued)

Table 22 (continued)

n	Decimation d	Polynomial $(c_0, c_1, c_2, c_3, c_4)$	Tap position $(r_1, r_2, r_3, r_4, r_5)$
	7	1 0 0 1 0	1 2 6 9 12
	7	1 0 0 1 0	1 3 10 12 13
	7	1 0 0 1 0	1 6 9 12 13
	7	1 0 1 0 0	3 5 7 8 9
	11	1 1 1 1 0	1 2 4 11 12
	11	1 1 1 1 0	1 2 9 10 11
	15	1 1 1 0 1	3 5 6 8 13
	15	1 1 1 1 0	3 5 7 8 9
15	1	1 1 1 0 1	4 5 12 13 14
	3	1 0 1 0 0	2 6 8 9 10
	3	1 0 1 0 0	4 5 6 7 14
	7	1 0 1 1 1	2 5 7 10 13
	7	1 0 1 1 1	2 5 8 11 14
	7	1 0 0 1 0	3 4 5 7 12
	11	1 0 0 1 0	2 3 6 7 13
	11	1 1 1 0 1	2 4 9 11 13
	11	1 0 1 1 1	2 9 10 11 12
	15	1 1 1 0 1	1 2 3 5 6
16	1	1 1 0 1 1	1 10 11 12 14
	1	1 1 1 0 1	1 10 11 12 14
	15	1 1 1 0 1	3 6 9 12 14
17	3	1 0 1 0 0	1 6 7 8 9
	3	1 1 0 1 1	4 7 8 9 12
	7	1 0 1 0 0	1 3 12 13 14
	7	1 1 0 1 1	1 4 10 11 13
	7	1 0 0 1 0	1 5 11 12 13
	11	1 1 1 0 1	1 3 6 12 13
	15	1 1 1 1 0	1 3 12 13 14
18	1	1 1 1 0 1	1 2 12 13 14
	3	1 1 0 1 1	4 7 8 10 15
	3	1 1 0 1 1	5 10 11 14 17
	7	1 0 0 1 0	1 2 5 7 11
	7	1 1 0 1 1	5 7 8 11 17
	11	1 0 0 1 0	1 8 9 11 15
	15	1 1 1 0 1	2 9 12 15 17
20	1	1 1 1 0 1	5 10 12 18 19

References

1. F.S. Annexstein, Generating de Bruijn sequences: an efficient implementation. IEEE Trans. Inf. Theory **46**(2), 198–200 (1997)
2. E.R. Berlekamp, *Algebraic Coding Theory*, Ch. 7 (McGraw-Hill, New York, 1968)
3. A.H. Chan, R.A. Games, E.L. Key, On the complexities of de Bruijn sequences. J. Combin. Theory Ser. A **33**(3) 233–246 (1982)
4. A.H. Chan, R.A. Games, J.J. Rushanan, On quadratic m-sequences, in *IEEE International Symposium on Information Theory*, vol. 364 (1994)
5. A.C. Chang, S.W. Golomb, G. Gong, P.V. Kumar, On the linear span of ideal autocorrelation sequences arising from the Segre hyperoval, in *Sequences and their Applications—Proceedings of SETA'98, Discrete Mathematics and Theoretical Computer Science* (Springer, London, 1999)
6. T. Chang, B. Park, Y.H. Kim, I. Song, An efficient implementation of the D-homomorphism for generation of de Bruijn sequences. IEEE Trans. Inf. Theory **45**(4), 1280–1283 (1999)
7. C. De Canniére, O. Dunkelman, M. Knežević, KATAN and KTANTAN—a family of small and efficient hardware-oriented block ciphers. in *Proceedings of the 11th International Workshop on Cryptographic Hardware and Embedded Systems*, LNCS, vol. 5747 (Springer, Heidelberg, 2009). pp. 272–288
8. J. Dillon, H. Dobbertin, New cyclic difference sets with singer parameters. Finite Fields Appl. **10**(3), 342–389 (2004)
9. H. Dobbertin, Kasami power functions, permutation polynomials and cyclic difference sets, in *Proceedings of the NATO-A.S.I. Workshop Difference Sets, Sequences and their Correlation Properties*, (Kluwer, Bad Windsheim/Dordrecht, 1999), pp. 133–158
10. E. Dubrova, A list of maximum period NLFSRs. Report 2012/166, Cryptology ePrint Archive (2012), http://eprint.iacr.org/2012/166.pdf
11. eSTREAM: The ECRYPT stream cipher project. http://www.ecrypt.eu.org/stream/
12. T. Etzion, A. Lempel, Construction of de Bruijn sequences of minimal complexity. IEEE Trans. Inf. Theory **30**(5), 705–709 (1984)
13. R. Evan, H.D.L. Hollman, C. Krattenthaler, Q. Xiang, Gauss sums, Jacobi sums and p-ranks of cyclic difference sets. J. Combin. Theory Ser. A, **87**(1), 74–119 (1999)
14. H. Fredricksen, A class of nonlinear de Bruijn cycles. J. Combin. Theory Ser. A **19**(2), 192–199 (1975)
15. H. Fredricksen, A survey of full length nonlinear shift register cycle algorithms. SIAM Rev. **24**(2), 195–221 (1982)
16. H. Fredricksen, I. Kessler, Lexicographic compositions and de Bruijn sequences. J. Combin. Theory Ser. A **22**, 17–30 (1977)
17. H. Fredricksen, J. Maiorana, Necklaces of beads in k colors and k-ary de Bruijn sequences. Discrete Math. **23**(3), 207–210 (1978)
18. R.A. Games, A generalized recursive construction for de Bruijn sequences. IEEE Trans. Inf. Theory **29**(6), 843–850 (1983)
19. B.M. Gammel, R. Göttfert, O. Kniffler, Achterbahn-128/80 (2006), http://www.ecrypt.eu.org/stream/p2ciphers/achterbahn/achterbahn_p2.pdf
20. S.W. Golomb, *Shift Register Sequences* (Aegean Park Press, Laguna Hills, 1981)
21. S.W. Golomb, On the classification of balanced binary sequences of period $2^n - 1$. IEEE Trans. Inf. Theory, 26(6), 730–732 (1980)
22. S.W. Golomb, G. Gong, *Signal Design for Good Correlation: for Wireless Communication, Cryptography, and Radar* (Cambridge University Press, New York, 2004)
23. G. Gong, Randomness and representation of span n sequences, in *Proceedings of the 2007 International Conference on Sequences, Subsequences, and Consequences, SSC'07* (Springer, Heidelberg, 2007), pp. 192–203
24. E.R. Hauge, T. Helleseth, De Bruijn sequences, irreducible codes and cyclotomy. Discrete Math. **159**(1–3), 143–154 (1996)

25. C.J.A. Jansen, W.G. Franx, D.E. Boekee, An efficient algorithm for the generation of de Bruijn cycles. IEEE Trans. Inf. Theory **37**(5), 1475–1478 (1991)
26. A. Lempel, On a homomorphism of the de Bruijn graph and its applications to the design of feedback shift registers. IEEE Trans. Comput. **C-19**(12), 1204–1209 (1970)
27. K. Mandal, Design and analysis of cryptographic pseudorandom number/sequence generators with applications in RFID. Ph.D. Thesis, University of Waterloo, 2013
28. K. Mandal, G. Gong, in *Cryptographically Strong de Bruijn Sequences with Large Periods*, ed. by L.R. Knudsen, H. Wu SAC 2012. LNCS, vol. 7707 (Springer, Heidelberg, 2012), pp. 104–118
29. K. Mandal, G. Gong, Cryptographic D-morphic analysis and fast implementations of composited De Bruijn sequences. Technical Report CACR 2012–27, University of Waterloo (2012)
30. K. Mandal, X. Fan, G. Gong, in *Warbler: A Lightweight Pseudorandom Number Generator for EPC Class 1 Gen 2 RFID Tags*, ed. by N.W. Lo, Y. Li. Cryptology and Information Security Series—The 2012 Workshop on RFID and IoT Security (RFIDsec'12 Asia), vol. 8 (IOS Press, Amsterdam, 2012), pp. 73–84
31. J.L. Massey, Shift-register synthesis and BCH decoding. IEEE Trans. Inf. Theory **15**(1), 122–127 (1969)
32. G.L. Mayhew, Weight class distributions of de Bruijn sequences. Discrete Math. **126**, 425–429 (1994)
33. G.L. Mayhew, Clues to the hidden nature of de Bruijn sequences. Comput. Math. Appl., **39**(11), 57–65 (2000)
34. G.L. Mayhew, S.W. Golomb, Linear Spans of modified de Bruijn sequences. IEEE Trans. Inf. Theory **36**(5), 1166–1167 (1990)
35. G.L. Mayhew, S.W. Golomb, Characterizations of generators for modified de Bruijn sequences. Adv. Appl. Math. **13**, 454–461 (1992)
36. J. Mykkeltveit, M.-K. Siu, P. Tong, On the cycle structure of some nonlinear shift register sequences. Inf. Control **43**(2), 202–215 (1979)
37. J.L.-F. Ng, Binary nonlinear feedback shift register sequence generator using the trace function, Master's Thesis, University of Waterloo, 2005
38. J.S. No, S.W. Golomb, G. Gong, H.K. Lee, P. Gaal, New binary pseudorandom sequences of period $2^n - 1$ with ideal autocorrelation. IEEE Trans. Inf. Theory **44**(2), 814–817 (1998)
39. T. Rachwalik, J. Szmidt, R. Wicik, J. Zablocki, Generation of nonlinear feedback shift registers with special-purpose hardware. Cryptology ePrint Archive, Report 2012/314 (2012), http://eprint.iacr.org/
40. J.-H. Yang, Z.-D. Dai, Construction of m-ary de Bruijn sequences (extended abstract), in *Advances in Cryptology—AUSCRYPT'92*, LNCS (Springer, Heidelberg, 1993), pp. 357–363