A Lower Bound for the Linear Span of Filtering Seqences

Charles C.Y. Lam[1] and Guang Gong[2]

[1] Department of Mathematics,
California State University, Bakersfield,
Bakersfield, California 93311, USA
clam@csub.edu
[2] Department of Electrical and Computer Engineering,
University of Waterloo,
Waterloo, Ontario N2L 3G1, CANADA
ggong@cacr.uwaterloo.ca

Abstract. Filtering generators are implemented in hardware to provide a simple and efficient method to generate sequences for use either as a key stream in stream cipher systems or as pseudorandom streams required in communications protocols. In 1985, Rueppel showed a lower bound for an equal-distanced tap positions case of the binary filtering generators. In this paper, we present an improved result by Gong which gives a lower bound on the linear span of a class of varied-distanced tap positions of binary filtering generators. The problem is eventually linked to a problem related to maximum distance separable (MDS) codes.

1 Introduction

Filtering generators [6, 17] have been studied for years since Key [10] and Groth's [8] work in the early 70's. Roughly speaking, a filtering generator can be described as follows (as in figure 1, we will give a formal definition later). Let f be a Boolean function with m variables, and n be the length of the linear feedback shift register (LFSR) which has a primitive connecting polynomial. The boolean function f which operates on m tap positions on the LFSR produces the filtering sequence. The problem of how to design such a generator with desired cryptographical properties have been slowly making progresses. One of the most important requirements in stream cipher design is to ensure the output of the key stream to possess a large linear span, in order to prevent the Berlekamp-Massey algorithm [1, 13] attack which reconstructs the entire sequence from a known portion of the key stream sequence. It is known that the linear span of a filtering sequence is bounded above from Key's discovery (for binary case, 1976 [10]) and Herlestam's result (for q-ary case, 1985 [9]). Rueppel, in his Ph.D thesis [18], observed a very nice sufficient condition such that the linear span of a filtering sequence is lower bounded by $\binom{n}{m}$ when the m tap position are equally distanced.

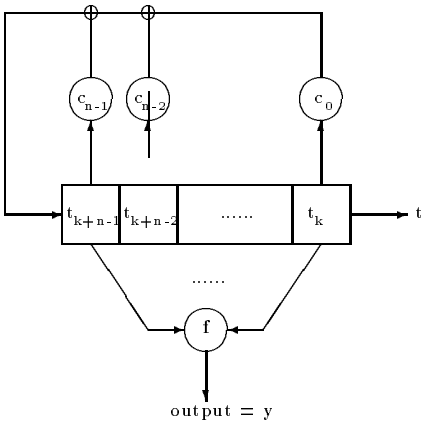

Fig. 1. A Construction of a Filter Function Generator

Massey and Serconek [14] introduced the notion of *generalised discrete Fourier transform*, and used it to re-prove Herlestam's result [9] and the Games-Chan algorithm [5] which provides a way to determine the linear span of a sequence of characteristic 2 with period 2^n.

On the other hand, there are surprising progresses on the cryptanalysis of filtering generators, namely, the algebraic attacks on filtering generators, proposed by Courtois and Meier in [4] in Eurocrypt 2003, and further investigated by Courtois in [3]. In the latter, the algebraic attack technique is applied to the stream cipher E_0 specified in Bluetooth Standard [2].

In 1990, Gong, in her Ph.D thesis [7], found a sufficient condition such that the linear span of the resulting filtering sequences are bounded below by the same bound as Rueppel's, but the m tap positions do not need to be equally distanced when some condition is satisfied. This result was never published. However, this is still a new result now. We feel it is necessary to let the community to know that there is another case for which the linear spans of the filtering sequences are bounded below. In this paper, we present this result in an improved version which is less constrained than the original version. The result is only dependent on one condition of the filter function f, the LFSR being used and its initial state. We consider this result to possibly provide some guidance in designing a good filtering generator. Another result of interest that comes with these conditions is that it is related to maximum distance separable (MDS) codes.

To conclude this section, we describe the trace representation of sequences defined over a general finite field $\mathbb{F}_q$.

Definition 1 *[16] Let $\mathbb{F}_{q^n}$ be an extension of $\mathbb{F}_q$. The* trace function $\mathrm{Tr}_q^{q^n}(\cdot)$ *maps an element from $\mathbb{F}_{q^n}$ to $\mathbb{F}_q$ with the formula*

$$\mathrm{Tr}_q^{q^n}(x) = x + x^q + x^{q^2} + \cdots + x^{q^{n-1}}.$$

We denote $\mathrm{Tr}_q^{q^n}(\cdot)$ *by* $\mathrm{Tr}(\cdot)$ *when the fields involved are understood from context. Note that the trace function is a linear function.*

Lemma 1 *[6] Let* $f(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0$ *be an irreducible polynomial over* $\mathbb{F}_q$ *of degree* n*, and* α *be a root of* $f(x)$*. Consider the sequence* $\underline{\mathbf{a}} = \{a_i\}_{i\geq 0}$ *defined by*

$$a_i = \mathrm{Tr}(\beta\alpha^i),\ i \geq 0,$$

for some $\beta \in \mathbb{F}_{q^n}$*. Then* $\underline{\mathbf{a}} \in G(f)$*. Note that* β *is not necessarily primitive.*

In order to present the new case that the linear spans of the filtering sequences are bounded below, we need to investigate the algebraic structures of these generators further. The mathematical background is presented in Section 2. In Section 3, we showed our improved result and provided a connection to maximum distance separable codes. An example of the analysis is illustrated in Section 4.

2 Algebraic Set Up of Filter Function Generators

Consider the set up as in figure 1. An LFSR of degree n with generator polynomial

$$h(t) = t^n + c_{n-1}t^{n-1} + c_{n-2}t^{n-2} + \cdots + c_1t + c_0 \in \mathbb{F}_2[t],$$

where $h(t)$ is primitive, is given as the feed for the filter function f. Suppose the tap positions are $l_1, l_2, \ldots, l_m$, then, at a particular stage of the LFSR with input bits $t_k, t_{k+1}, \ldots, t_{k+n-2}, t_{k+n-1}$, the function f takes in the m bits $t_{k+l_1}, t_{k+l_2}, \ldots, t_{k+l_m}$ out of the n bits of the LFSR, and outputs one bit y. As the LFSR clock proceeds, the output of f generates the filtering sequence, called the output sequence $\mathbf{y}$. The filter function f is added in an attempt to increase the linear span of the output sequence $\mathbf{y}$.

The filtering function f is of the form

$$f : \mathbb{F}_2^m \to \mathbb{F}_2.$$

To any basis $B = \{\beta_0, \beta_1, \ldots, \beta_{m-1}\} \subset \mathbb{F}_{2^m}$ over $\mathbb{F}_2$, we can associate each m-tuple over $\mathbb{F}_2^m$ with an element in $\mathbb{F}_{2^m}$ through a bijective mapping $\phi : \mathbb{F}_2^m \to \mathbb{F}_{2^m}$ where

$$\phi(z_0, z_1, \ldots, z_{m-1}) = \sum_{i=0}^{m-1} z_i\beta_i.$$

A boolean function f in m variables can be associated with a polynomial function F over $\mathbb{F}_{2^m}$ through ϕ such that

$$f(z_0, z_1, \ldots, z_{m-1}) = F(\sum_{i=0}^{m-1} z_i\beta_i).$$

Such a polynomial function F can be obtained from the boolean function f through the discrete Fourier transform or interpolation.

Let $q = 2^m$, and let $\{x_0, x_1, \ldots x_{q-1}\}$ be the collection of all elements of $\mathbb{F}_{2^m}$, with $x_0 = 0$. Let

$$F(x_i) = v_i, \ i = 0, 1, \ldots, q-1$$

where $v_i \in \mathbb{F}_2$ and consider the matrix

$$M = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_1 & x_2 & \cdots & x_{q-1} \\ 0 & x_1^2 & x_2^2 & \cdots & x_{q-1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_1^{q-1} & x_2^{q-1} & \cdots & x_{q-1}^{q-1} \end{bmatrix}.$$

By interpolation, we can obtain a Mattson-Solomon polynomial [15] function G of degree $q-1$ such that

$$G(x) = F(x), \ \forall x \in \mathbb{F}_q.$$

Let

$$G(x) = \sum_{i=0}^{q-1} d_i x^i,$$

then we have

$$(d_0, d_1, \ldots, d_{q-1})M = (v_0, v_1, \ldots, v_{q-1}).$$

Consider the matrix

$$B = \begin{bmatrix} 1 & 0 & 0 & \cdots & 1 \\ 0 & x_1^{-1} & x_1^{-2} & \cdots & x_1^{-(q-1)} \\ 0 & x_2^{-1} & x_2^{-2} & \cdots & x_2^{-(q-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_{q-1}^{-1} & x_{q-1}^{-2} & \cdots & x_{q-1}^{-(q-1)} \end{bmatrix}.$$

We will show that $B = M^{-1}$. Let $Q = (Q_{i,j}) = MB$. First, we need the following lemma:

Lemma 2

$$\sum_{x \in \mathbb{F}_q, x \neq 0} x^c = \begin{cases} 0, c \not\equiv 0 \pmod{q-1} \\ 1, c \equiv 0 \pmod{q-1}. \end{cases}$$

Proof: Consider

$$0 = x^{q-1} - 1 = (x-1)(x^{q-2} + \cdots + x + 1).$$

Let α be a primitive element in $\mathbb{F}_q$. If $c \not\equiv 0 \pmod{q-1}$, then $\alpha^c - 1 \neq 0$. Hence

$$(\alpha^{q-2})^c + (\alpha^{q-3})^c + \cdots + \alpha^c + 1 = 0,$$

but this is equivalent to

$$\sum_{x \in \mathbb{F}_q, x \neq 0} x^c = 0.$$

When $c \equiv 0 \pmod{q-1}$, $\alpha^c = 1$. Hence

$$\sum_{x \in \mathbb{F}_q, x \neq 0} x^c = \sum_{x \in \mathbb{F}_q, x \neq 0} 1 = 1.$$

□

Now, we proceed to compute the entries of $Q_{i,j}$ for $0 \leq i, j \leq q-1$. It can be seen easily that $Q_{0,0} = 1$ and $Q_{i,0} = 0$ for $0 < i \leq q-1$. For $0 < j < q-1$,

$$Q_{0,j} = \sum_{i=1}^{q-1} x_i^{-j} = \sum_{x \in \mathbb{F}_q, x \neq 0} x^{-j} = 0,$$

and for $j = q-1$,

$$Q_{0,q-1} = 1 + \sum_{i=1}^{q-1} x_i^{-(q-1)} = 0.$$

Also, when $0 < i, j \leq q-1$,

$$\begin{aligned} Q_{i,j} &= \sum_{k=0}^{q-1} M_{i,k} B_{k,j} = M_{i,0} B_{0,j} + \sum_{k=1}^{q-1} M_{i,k} B_{k,j} \\ &= 0 + \sum_{k=1}^{q-1} x_k^{i-j} = \sum_{t \in \mathbb{F}_q, t \neq 0} t^{i-j} = \begin{cases} 1, i = j \\ 0, i \neq j. \end{cases} \end{aligned}$$

Therefore, $MB = I$, and hence $B = M^{-1}$. We have

$$(d_0, d_1, \ldots, d_{q-1}) = (v_0, v_1, \ldots, v_{q-1}) M^{-1},$$

which gives

$$\begin{aligned} d_0 &= v_0 = F(0) \\ d_i &= \sum_{j=1}^{q-1} v_j x_j^{-i} = \sum_{t \in \mathbb{F}_q, t \neq 0} F(t) t^{-i} \text{ for } 1 \leq i \leq q-2 \\ d_{q-1} &= F(0) + \sum_{t \in \mathbb{F}_q, t \neq 0} F(t) t^{-(q-1)} = \sum_{t \in \mathbb{F}_q} F(t). \end{aligned}$$

Therefore, the coefficients of the boolean function in polynomial form can be calculated completely using interpolation with the above formula.

3 A Lower Bound for the Filtering Sequence

In this section, we combine the feeding LFSR and the filter function to evaluate the properties of the combined function. A lower bound on the linear complexity is given, provided the two functions satisfy certain conditions.

Let $\mathbf{y} = \{y_0, y_1, \ldots\}$ be the output of the filter function F. First consider the simple case $F(x) = x^r$, where $r = 2^{i_1} + 2^{i_2} + \cdots + 2^{i_s} \leq q-1$ for some distinct $i_1, i_2, \ldots, i_s \in \mathbb{Z}$ with $0 \leq i_1, i_2, \ldots, i_s \leq m-1$. Pick a basis $\beta_1, \beta_2, \ldots, \beta_m \in \mathbb{F}_{2^m}$ over $\mathbb{F}_2$. Suppose that at stage j of the LFSR, we have input bits $t_j, t_{j+1}, \ldots, t_{j+n-1}$ and tap positions $l_1, l_2, \ldots, l_m$, then the bits $t_{j+l_1}, t_{j+l_2}, \ldots, t_{j+l_m}$ are taken as input to the filter function. For simplicity, we rename $t_{j+l_r} = t_{j_r}, 1 \leq r \leq m$. That is, the bits $t_{j_1}, t_{j_2}, \ldots, t_{j_m}$ are chosen as input to the filter function.

We consider each input x_j to the filter function as an element over $\mathbb{F}_{2^m}$ by

$$x_j = t_{j_1}\beta_1 + t_{j_2}\beta_2 + \cdots + t_{j_m}\beta_m, \ j \geq 0.$$

Let $\mathbf{x} = \{x_0, x_1, \ldots\}$ be the sequence over $\mathbb{F}_q$ that is fed to the filter function F, we denote $\mathbf{x}^r = \{x_0^r, x_1^r, \ldots\}$.

Lemma 3 *Suppose $h(t) \in \mathbb{F}_2[t]$ is of degree n. If $\mathbf{x} \in G_{\mathbb{F}_q}(h)$, then $\mathbf{x}^{2^i} \in G_{\mathbb{F}_q}(h)$.*

Proof: Let $h(t) = t^n - c_{n-1}t^{n-1} - \cdots - c_0$, where $c_i \in \mathbb{F}_2$. Then, at stage j, the bits $t_{j_1}, t_{j_2}, \ldots, t_{j_m}$ form the following recurrence relations

$$\begin{aligned}
t_{j_1+n} &= c_{n-1}t_{j_1+n-1} + c_{n-2}t_{j_1+n-2} + \cdots + c_0 t_{j_1}, \\
t_{j_2+n} &= c_{n-1}t_{j_2+n-1} + c_{n-2}t_{j_2+n-2} + \cdots + c_0 t_{j_2}, \\
\cdots &= \cdots \\
t_{j_m+n} &= c_{n-1}t_{j_m+n-1} + c_{n-2}t_{j_m+n-2} + \cdots + c_0 t_{j_m}.
\end{aligned}$$

By multiplying $\beta_1, \beta_2, \cdots, \beta_m$ to the respective equations above, and take the sum of the equations, we see that

$$x_k = c_{n-1}x_{k-1} + c_{n-2}x_{k-2} + \cdots + c_0 x_{k-n}.$$

Hence

$$\begin{aligned}
x_k^{2^i} &= c_{n-1}^{2^i}x_{k-1}^{2^i} + c_{n-2}^{2^i}x_{k-2}^{2^i} + \cdots + c_0^{2^i}x_{k-n}^{2^i} \\
&= c_{n-1}x_{k-1}^{2^i} + c_{n-2}x_{k-2}^{2^i} + \cdots + c_0 x_{k-n}^{2^i}.
\end{aligned}$$

□

Lemma 4 *Let $\gcd(m, n) = 1$, and α be a primitive root of an irreducible polynomial $h(x) \in \mathbb{F}_2[x]$ of degree n. Then*

$$\{\alpha^{2^i} \mid i = 0, 1, \ldots n-1\} = \{\alpha^{2^{mi}} \mid i = 0, 1, \ldots n-1\}$$

Let $\alpha^{2^{ms_i}} = \alpha^{2^i}$ for $0 \leq i \leq n-1$. Then

$$ms_i \equiv i \pmod{n}.$$

Proof: Since $h(x)$ is irreducible over $\mathbb{F}_2$, all the roots of h are

$$T_1 = \{\alpha^{2^i} \mid i = 0, 1, \ldots n-1\}.$$

Since $\gcd(m, n) = 1$, h is also an irreducible polynomial over $\mathbb{F}_{2^m}$. Therefore the roots of h are

$$T_2 = \{(\alpha^{2^m})^i \mid i = 0, 1, \ldots n-1\} = \{(\alpha^{2^{mi}}) \mid i = 0, 1, \ldots n-1\}.$$

We must have $T_1 = T_2$. Hence for every i, there exists a permutation $s(i) = s_i$ such that

$$\alpha^{2^{ms_i}} = \alpha^{2^i}.$$

Since $\alpha^{2^n} = \alpha$ and $\gcd(2^{ms_i}, 2^n - 1) = 1$, we must have

$$ms_i \equiv i \pmod{n}.$$

□

Given h, by Lemma 1, each element $x_k \in \mathbf{x}$ can be written as

$$x_k = \sum_{i=0}^{n-1} A_i \alpha_i^k$$

where $\alpha_i = \alpha^{2^{mi}}$ and $A_i \in \mathbb{F}_{q^n}$. Hence

$$\begin{aligned} x_k^{2^j} &= \sum_{i=0}^{n-1} A_i^{2^j} \alpha_i^{2^j k} = \sum_{i=0}^{n-1} A_{ij} \alpha^{2^{mi+j}k} \text{ where } A_{ij} = A_i^{2^j} \\ &= \sum_{i=0}^{n-1} A_{s_i j} \alpha^{2^{ms_i+j}k} = \sum_{i=0}^{n-1} A_{s_i j} \alpha^{2^{i+j}k} = \sum_{i=0}^{n-1} C_{ij} \alpha^{2^i k}, \end{aligned}$$

where $C_{ij} = A_{s_{i-j(mod n)}, j}$. Hence

$$\begin{aligned} x_k^r = x_k^{2^{i_1}+2^{i_2}+\cdots+2^{i_s}} &= \prod_{t \in \{i_1, i_2, \ldots, i_s\}} \left(\sum_{i=0}^{n-1} C_{it} \alpha^{2^i k} \right) \\ &= \sum_{1 \le w(v) \le r} B_{v,r} \alpha^{vk}, \end{aligned}$$

where $w(v)$ is the hamming weight of v, and $B_{v,r} \in \mathbb{F}_{q^n}$.

Now we consider the general form of the function F. Let $y_k = F(x_k)$. Then

$$\begin{aligned} y_k = F(x_k) &= \sum_{i=0}^{2^m-1} d_i x_k^i = \sum_{i=0}^{2^m-1} \left(d_i \sum_{1 \le w(v) \le w(i)} B_{v,i} \alpha^{vk} \right) \\ &= \sum_{1 \le w(v) \le m} \left(\sum_{i=0}^{2^m-1} d_i B_{v,i} \right) \alpha^{vk}. \end{aligned}$$

Substitute

$$B_v = \sum_{i=0}^{2^m-1} d_i B_{v,i},$$

then

$$y_k = \sum_{1 \le w(v) \le m} B_v \alpha^{vk}.$$

Lemma 5 *$A_{ij} = \omega^{2^{mi+j}}$ for some $\omega \in \mathbb{F}_{2^{mn}} = \mathbb{F}_{q^n}$.*

Proof: Since $\mathbf{x} \in G_{\mathbb{F}_{2^m}}(h)$, by Lemma 1, there exists $\omega \in \mathbb{F}_{2^{mn}}$ such that

$$x_k = \mathrm{Tr}_{2^m}^{2^{mn}}(\omega \alpha^k) = \sum_{i=0}^{n-1} \omega^{2^{mi}} \alpha^{2^{mi}k} = \sum_{i=0}^{n-1} \omega^{2^{ms_i}} \alpha^{2^{ms_i}k} = \sum_{i=0}^{n-1} \omega^{2^{ms_i}} \alpha^{2^i k}.$$

Hence

$$x_k^{2^j} = \sum_{i=0}^{n-1} \omega^{2^{ms_i+j}} \alpha^{2^{i+j}k}.$$

However

$$x_k^{2^j} = \sum_{i=0}^{n-1} A_{s_i j} \alpha^{2^{i+j}k},$$

therefore

$$A_{s_i j} = \omega^{2^{ms_i+j}}.$$

□

Using the above fact, we can write the system as

$$\begin{bmatrix} x_k \\ x_k^2 \\ \vdots \\ x_k^{2^{m-1}} \end{bmatrix} = \begin{bmatrix} \omega^{2^{ms_0}} & \omega^{2^{ms_1}} & \cdots & \omega^{2^{ms_{n-1}}} \\ \omega^{2^{ms_{n-1}+1}} & \omega^{2^{ms_0+1}} & \cdots & \omega^{2^{ms_{n-2}+1}} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{2^{ms_{n-(m-1)}+m-1}} & \omega^{2^{ms_{n-(m-2)}+m-1}} & \cdots & \omega^{2^{ms_{n-m}+m-1}} \end{bmatrix} \begin{bmatrix} \alpha^k \\ \alpha^{2k} \\ \vdots \\ \alpha^{2^{n-1}k} \end{bmatrix}. \tag{1}$$

We denote the $m \times n$ coefficient matrix to the α's above as W.

On the other hand, we have

$$\begin{aligned} y_k &= \sum_{i=0}^{2^m-1} \sum_{1 \le w(v) \le m} d_i B_{v,i} \alpha^{vk} \\ &= \sum_{i=0}^{2^m-2} \sum_{1 \le w(v) < w(i) < m} d_i B_{v,i} \alpha^{vk} \; + \; d_{2^m-1} \sum_{1 \le w(v) \le m} B_{v,2^m-1} \alpha^{vk}, \end{aligned}$$

since the only element of weight m between 0 and $2^m - 1$ is $2^m - 1$. If $w(v) = m$, then α^{vk} appears only in the second half of the expression. Also, note that since

$$y_k = F(x_k) = \sum_{i=0}^{2^m-1} x_k^i,$$

α^{vk} can only appear in the spectrum of x^{2^m-1}. Hence if $d_{2^m-1} \neq 0$, then, by counting the number of nonzeros of the coefficient of α^{vk} for each v with $w(v) = m$, we can find a lower bound for the linear span of the output sequence $\mathbf{y}$.

Let $w(v) = m$. Then $v = 2^{u_1} + 2^{u_2} + \cdots + 2^{u_m}$ for some $u_1, u_2, \ldots, u_m \in \mathbb{Z}$. Let $\tau = \{u_1, u_2, \ldots, u_m\}$. Denote the rows and columns of W by $\{1, 2, \ldots, n\}$ and $\{1, 2, \ldots, m\}$ respectively, we get

$$\begin{aligned} B_{v,2^m-1} = [\alpha^{vk}]\ & (\omega^{2^{ms_0}}\alpha^k + \omega^{2^{ms_1}}\alpha^{2k} + \cdots + \omega^{2^{ms_{n-1}}}\alpha^{2^{n-1}k}) \\ & \cdot(\omega^{2^{ms_0+1}}\alpha^{2k} + \omega^{2^{ms_1+1}}\alpha^{2^2k} + \cdots + \omega^{2^{ms_{n-1}+1}}\alpha^{2^nk}) \\ & \cdots(\omega^{2^{ms_0+m-1}}\alpha^{2^{m-1}k} + \cdots + \omega^{2^{ms_{n-1}+m-1}}\alpha^{2^{m-1+n-1}k}), \end{aligned}$$

which is also the determinant of the matrix W_τ formed by the columns $u_1 + 1, u_2 + 1, \ldots, u_m + 1$ of W. Hence if for every v with $w(v) = m$ such that $v = 2^{u_{v_1}} + 2^{u_{v_2}} + \cdots + 2^{u_{v_m}}$ and $\tau_v = \{u_{v_1}, u_{v_2}, \ldots, u_{v_m}\}$, we have $\det(W_{\tau_v}) \neq 0$, then

$$\mathrm{LS}(\mathbf{y}) \geq \binom{n}{m}.$$

Lemma 6 *Let $b = 2^n$, and $\omega^{2^{ms_j}} = \omega_j$. By rearranging the rows of W, we can get the matrix*

$$P_W = \begin{bmatrix} \omega_0 & \omega_1 & \cdots & \omega_{n-1} \\ \omega_0^b & \omega_1^b & \cdots & \omega_{n-1}^b \\ \vdots & \vdots & \ddots & \vdots \\ \omega_0^{b^{m-1}} & \omega_1^{b^{m-1}} & \cdots & \omega_{n-1}^{b^{m-1}} \end{bmatrix} \tag{2}$$

Proof: It can be seen from the matrix W that

$$W_{i,j} = \omega^{2^{ms_{(j-i) \pmod n}+i-1}},\ 1 \leq i \leq m,\ 1 \leq j \leq n.$$

Let $W_{i,j} = \omega^{2^{p_{i,j}}}$. First notice that the powers $p_{i,j}$ form the set $\mathbb{Z}_{mn}$ such that each $p_{i,j}$ is distinct. Then, we show that for any fixed $1 \leq i_1, i_2 \leq m$ and $1 \leq j_1, j_2 \leq n$, we must have

$$p_{i_1,j_1} - p_{i_2,j_1} = p_{i_1,j_2} - p_{i_2,j_2} \pmod{mn}. \tag{3}$$

Observe that

$$s_i \equiv ik \pmod n$$

for some $k \in \mathbb{Z}$. Hence

$$\begin{aligned} p_{i_1,j_1} - p_{i_2,j_1} &\equiv ms_{(j_1-i_i) \pmod n} + i_i - 1 + ms_{(j_1-i_2) \pmod n} + i_2 - 1 \pmod{mn} \\ &\equiv m(s_{(j_1-i_i) \pmod n} - s_{(j_1-i_2) \pmod n}) + i_i - 1 + i_2 - 1 \pmod{mn} \\ &\equiv m((j_1 - i_i)k - (j_1 - i_2)k) + i_i - 1 + i_2 - 1 \pmod{mn} \\ &\equiv m((j_2 - i_i)k - (j_2 - i_2)k) + i_i - 1 + i_2 - 1 \pmod{mn} \\ &\equiv p_{i_1,j_2} - p_{i_2,j_2} \pmod{mn}. \end{aligned}$$

Then, we show that for any $1 \leq i_1, i_2 \leq m$, in the first column of the matrix W,

$$p_{i_1,1} - p_{i_2,1} \equiv 0 \pmod{n}. \tag{4}$$

We have

$$\begin{aligned} p_{i_1,1} - p_{i_2,1} &\equiv ms_{(1-i_1) \pmod n} + i_1 - 1 - (ms_{(1-i_2) \pmod n} + i_2 - 1) \pmod n \\ &\equiv (1-i_1) - (1-i_2) + i_1 - 1 - (i_2 - 1) \pmod n \\ &\equiv 0 \pmod n. \end{aligned}$$

Now with (3) and (4), and that $p_{1,1} = 0$ and the set $\{p_{1,1}, p_{2,1}, \ldots p_{m,1}\}$ is a set of m distinct elements in $\mathbb{Z}_{mn}$, with $0 \leq p_{i,j} = ms_{(j-i) \pmod n} \leq mn - 1$, we must have

$$\{p_{1,1}, p_{2,1}, \ldots, p_{m,1}\} = \{0, n, 2n, \ldots, (m-1)n\}.$$

In other words, a rearrangement of the rows of W gives

$$[\omega^{2^0}, \omega^{2^n}, \omega^{2^{2n}}, \ldots, \omega^{2^{(m-1)n}}]$$

as the first column of the matrix. Substitute $b = 2^n$, $\omega^{2^{ms_j}} = \omega_j$, and the result follows. □

Lemma 7 *[11, Lemma 3.51] Let $\beta_1, \beta_2, \ldots \beta_m$ be elements of $\mathbb{F}_{q^n}$, then the determinant*

$$\begin{vmatrix} \beta_1 & \beta_2 & \cdots & \beta_m \\ \beta_1^q & \beta_2^q & \cdots & \beta_m^q \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{q^{m-1}} & \beta_2^{q^{m-1}} & \cdots & \beta_m^{q^{m-1}} \end{vmatrix}$$

is nonzero if and only if $\beta_1, \beta_2, \ldots \beta_m$ are linearly independent over $\mathbb{F}_q$.

Using the above lemma, we can define a condition for the output sequence to have a guaranteed linear span.

Theorem 1 *Consider the filtering generator system with an LFSR of degree n and a filter function with m tap positions, where $\gcd(m, n) = 1$, $d_{2^m - 1} \neq 0$. Let $q = 2^m$ and $\tau = \{\omega, \omega^q, \ldots, \omega^{q^{n-1}}\}$ where $\omega \in \mathbb{F}_{2^{mn}}$, as defined in the trace representation of the output sequence from the LFSR feeding into the filter function. Then*

$$\mathrm{LS}(\mathbf{y}) \geq \binom{n}{m}$$

if and only if any m elements in τ are linearly independent over $\mathbb{F}_{2^n}$.

Proof: Consider any m columns $\{i_1, i_2, \ldots, i_m\}$ of P_W, and let $\beta_j = \omega_{i_j}$ for $1 \leq j \leq m$. Applying $q = 2^n$ in lemma 7, the result follows. □

If the conditions in Theorem 1 for ω cannot be completely satisfied, we can still give a lower bound for the sequence as follows.

Corollary 1 *Using the notations as above, let μ be the number of m subsets of τ that are linearly independent over F_{2^n}. Then*

$$\mathrm{LS}(\mathbf{y}) \geq \mu.$$

Remark: Note that the requirement that $d_{2^m-1} \neq 0$ implies that the output sequence $\mathbf{y}$ is not balanced. However it is easy to design a function f such that $|N_1 - N_0| = 2$, where N_i represents the number of occurances of the symbol i for all inputs to f, by changing one value of the output of a balanced function g. The predictability of the filtering sequence is compromised in that

$$|\Pr(y_i = 0) - \Pr(y_i = 1)| = \frac{1}{2^{m-1}}.$$

However, since the advantage of predicting the output is still of exponential order in m, we do not consider the difference as significant when a large m is used.

Remark: Note also that we only require the filter function to possess the property that $d_{2^m-1} \neq 0$. Otherwise, the linear span property depends only on the LFSR used and its initial state, which in turn determines ω.

Remark: In [7], Gong showed the following.

Fact 1 *Consider the filtering generator system with the parameters m, n, where $\gcd(m, n) = 1$, $n \equiv -1 \pmod{m}$, $d_{2^m-1} \neq 0$. Let $q = 2^m$ and $\tau = \{\omega, \omega^q, \ldots, \omega^{q^{n-1}}\}$ where $\omega \in \mathbb{F}_{2^{mn}}$. Then*

$$\mathrm{LS}(\mathbf{y}) \geq \binom{n}{m}$$

if and only if any m elements in τ are linearly independent over $\mathbb{F}_{2^n}$.

The improved result (Theorem 1) removes the requirement that $n \equiv -1 \pmod{m}$.

The results we obtained also imply a construction for an MDS algebraic error-correcting code [12]. Consider the element $\omega \in \mathbb{F}_{2^{mn}}$ such that the set $\tau = \{\omega, \omega^q, \ldots, \omega^{q^{n-1}}\}$ satisfies Theorem 1. Let $\gamma_0, \gamma_1, \ldots, \gamma_{m-1}$ be a basis for $\mathbb{F}_{2^{mn}}$ over $\mathbb{F}_{2^n}$. We can express each element $\omega^{q^i}, 0 \leq i \leq n-1$ as

$$\omega^{q^i} = \sum_{j=0}^{m-1} v_{j,i}\gamma_j,$$

where $v_{j,i} \in \mathbb{F}_{2^n}$.

Theorem 2 *The condition that any m elements in τ are linearly independent is equivalent to that any m columns are linearly independent over $\mathbb{F}_{2^n}$ in the matrix*

$$C = \begin{bmatrix} v_{0,0} & v_{0,1} & \cdots & v_{0,n-1} \\ v_{1,0} & v_{1,1} & \cdots & v_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ v_{m-1,0} & v_{m-1,1} & \cdots & v_{m-1,n-1} \end{bmatrix},$$

where C is a matrix over $\mathbb{F}_{2^n}$, and can also be considered as the parity check matrix of an MDS code of dimension m over $\mathbb{F}_{2^n}$.

There are currently no theoretical results on the conditions on ω to generate a set τ that gives the condition in Theorem 1. One direction of further research is to see if there is a class of known MDS codes over fields of characteristic 2 that can generate such an element ω.

Open Problem 1 *Find an efficient generating method for $\omega \in \mathbb{F}_{2^{mn}}$ such that any m elements in $\tau = \{\omega, \omega^q, \ldots, \omega^{q^{n-1}}\}$ are linearly independent over $\mathbb{F}_{2^n}$.*

4 An Example

Consider the filter function generator with an LFSR of length 5 defined by

$$h(t) = t^5 + t^2 + 1.$$

Filter function f is defined by taking the first, second, and fifth bit such that

$$f(t_1, t_2, t_5) = t_2 + t_1\bar{t_2}\bar{t_5}.$$

The hardware implementation of this function is illustrated in figure 2.

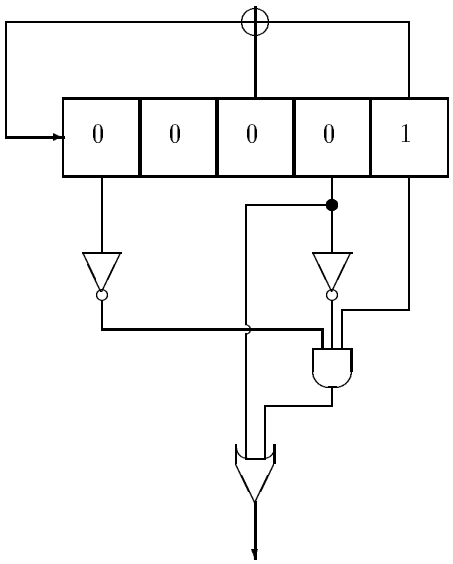

Fig. 2. Numerical Example of a Filter Function Generator

The output of the LFSR generates the sequence

$$1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, \ldots$$

with a linear span of 5. The output of the filter function hence generates the sequence

$$\mathbf{y} = 1,0,0,0,0,1,0,0,1,1,1,1,0,0,1,1,1,1,1,0,0,0,1,1,1,1,1,1,1,1,1,1,0,0,\ldots$$

It can be shown using the Berlekamp-Massey Algorithm that the linear span of this output sequence is 25, with generator polynomial

$$g(t) = t^{25} + t^{24} + t^{21} + t^{19} + t^{18} + t^{16} + t^{15} + t^{14} + t^{13} + t^{11} + t^9 + t^5 + t^2 + t + 1.$$

Note that 25 is also the maximum linear span that this configuration can achieve, since, by Herlestam's result,

$$\mathrm{LS}(\mathbf{y}) \leq \sum_{i=1}^{3} \binom{5}{i} = 25.$$

Let $\mathbb{F}_{2^{15}}$ be the field $\mathbb{F}_2[z]/(z^{15}+z+1)$. Let $\alpha = z$ be the generating element. Then

$$\beta = \alpha^{1+2^3+2^6+2^9+2^{12}}$$

is a generating element for the subfield $\mathbb{K}_3 = \mathbb{F}_{2^3}$, and

$$\gamma = \alpha^{1+2^5+2^{10}}$$

is a generating element for the subfield $\mathbb{K}_5 = \mathbb{F}_{2^5}$.

It can be shown that $\{\beta, \beta^2, \beta^3\}$ is a basis for $\mathbb{K}_3$ over $\mathbb{F}_2$. Hence for each input (t_1, t_2, t_5), we can associate the input by the finite field element in $\mathbb{K}_3$ with

$$\phi(t_1, t_2, t_5) = x = t_1\beta + t_2\beta^2 + t_5\beta^3.$$

Using the method discussed in section 2, we can convert the boolean function f into the polynomial form F such that with the mapping ϕ,

$$F(x) = x^7 + \beta^2 x^6 + \beta^4 x^5 + \beta^4 x^4 + \beta x^3 + \beta^2 x^2 + \beta x.$$

Note that the value d_7, which is the coefficient to x^7, is 1. Hence we can apply Theorem 1.

The input of the sequence to the filter function, using the form over the finite field, is the following sequence.

$$\begin{aligned}\mathbf{x} = {} & \beta, \beta^3, 0, 0, \beta^5, \beta, \beta^3, \beta^5, \beta, \beta^2, \beta^6, 1, \beta^3, \beta^5, \beta^6, \beta^4, \\ & \beta^4, \beta^4, 1, \beta^3, 0, \beta^5, \beta^6, \beta^7, \beta^2, \beta^6, \beta^4, 1, \beta^2, \beta, \beta^2, \beta, \beta^3, \ldots\end{aligned}$$

Using the discrete Fourier transform, we get

$$\mathbf{x}_k = \mathrm{Tr}_{2^3}^{2^{15}}(\alpha^{29009}\gamma^{-11k}).$$

Let $\omega = \alpha^{29009}$. It can be shown that ω satisfies the conditions in Theorem 1, with $n = 5$, $m = 3$, and $d_7 = 1$. Hence the output sequence from the filter function has linear span

$$\mathrm{LS}(\mathbf{y}) \geq \binom{5}{3} = 10.$$

References

1. E.R. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill, New York, 1968.
2. Specification of the Bluetooth system, available at `www.bluetooth.com`
3. N. Courtois, Fast algebraic attacks on stream ciphers with linear feedback, *Advances in Cryptology – Crypto'2003*, Lecture Notes in Computer Science, 2729, pp. 176-194, Springer-Verlag, 2003.
4. N. Courtois and W. Meier, Algebraic attacks on stream ciphers with lienar feedback, *Advances in Cryptology – EuroCrypt'2003*, Lecture Notes in Computer Science, 2656, pp. 345-359, Springer-Verlag, 2003.
5. R.A. Games and A.H. Chan, A Fast Algorithm for Determining the Complexity of a Binary Sequence with Period 2, *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 144-146, Jan 1983.
6. S.W. Golomb, *Shift Register Sequences,* Holden-Day, Inc., 1967. Revised edition, Aegean Park Press, 1982.
7. G. Gong, Analysis and Synthesis of Phases and Linear Complexity of Non-Linear Feedforward Sequences, *Ph.D. thesis*, University of Elec. Sci. and Tech. of China, 1990.
8. E.J. Groth, Generation of Binary Sequences with Controllable Complexity, IEEE Trans. Info. Theory, Vol IT-17, No.3, pp. 288-296, May 1971.
9. T. Harlestam, On Functions of Linear Shift Register Sequences, *Advances in Cryptology – Eurocrypt 85'*, Lecture Notes in Computer Science, LNCS 0219, pp. 119-129, Springer-Verlag, 1985.
10. E. Key, An Analysis of the Structure and Complexity of Nonlinear Binary Sequence Generators, *IEEE Trans. Information Theory*, 22, pp. 732-736, 1976.
11. R. Lidl and H. Niederreiter, *Introduction to Finite Fields and Their Applications*, Cambridge University Press, 1994.
12. F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland Publishing Company, 1977.
13. J. L. Massey, Shift-register Synthesis and BCH Decoding, *IEEE Trans. Information Theory*, Vol. 15, No. 1, pp. 122-127, January 1969.
14. J. L. Massey, S. Serconek, Linear Complexity of Periodic Sequences: A General Theory, *Advances in Cryptology – Crypto 96'*, Lecture Notes in Computer Science, LNCS 1109, pp. 358-371, Springer-Verlag, 1996.
15. H.F. Mattson and G. Solomon, A New Treatment of Bose-Chaudhuri Codes, *J. SIAM*, vol. 9, pp. 654-600, December 1961.
16. A.J. Menezes (editor), *Applications of Finite Fields*, Kluwer Academic Publishers, 1993.
17. A.J. Menezes, P.C. van Oorschot, S.A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1997.
18. R.A. Rueppel, *Analysis and Design of Stream Ciphers*, Springer-Verlag, 1986.