Cascaded GMW Sequences*

Andrew Klapper† A. H. Chan ‡ Mark Goresky§

November 17, 2002

Abstract

Pseudorandom binary sequences with high linear complexity and low correlation function values are sought in many applications of modern communication systems. A new family of pseudorandom binary sequences, cascaded GMW sequences, is constructed. These sequences are shown to share many desirable correlation properties with the GMW sequences of Gordon, Mills, and Welch, for example high shifted autocorrelation values and, in many cases, three valued cross-correlation values with m-sequences. It is shown, moreover, that in many cases the linear complexities of cascaded GMW sequences are far greater than those of GMW sequences.

Keywords: Pseudorandom sequences, linear complexity, cross correlations, cryptography, GMW sequences.

I Introduction

Pseudorandom binary sequences with high linear complexity and low correlation function values are sought in many applications of modern communication systems. The class of binary "geometric sequences" based on odd primes [4] has recently attracted considerable attention due to their ease of generation using shift register hardware, and their enormous

*Parts of this work have been presented at the Twenty-Eighth Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, October 1990. Work by the first and second authors sponsored in part by the National Security Agency under Grant Number MDA904-91-H-0012. The United States Government is authorized to reproduce and distribute reprints notwithstanding any copyright notation hereon.

†The University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada, and Northeastern University, Boston, MA 02115.

‡Northeastern University, Boston, MA 02115.

§Northeastern University, Boston, MA 02115.

linear complexity. However it was shown recently [10] that the periodic autocorrelation function values of the geometric sequences (based on odd primes) is unacceptably high. On the other hand, the binary sequences of Gordon, Mills and Welch [8, 17] have the same autocorrelation function as an m-sequence of the same period. Although the linear complexity of a GMW sequence is much greater than that of an m-sequence of the same period, it is desirable to achieve higher linear complexities. This would provide higher levels of cryptographic security against attacks using the Berlekamp-Massey algorithm [14] (recall that the Berlekamp-Massey algorithm can be used to determine a sequence from a short subsequence when the sequence has low linear complexity).

In this paper we construct a new family of pseudorandom binary sequences, cascaded GMW sequences, which share many of the most desirable properties of both the geometric sequences and the GMW sequences. They have the same periodic autocorrelation function values as m-sequences of the same period, but in many cases have greater linear complexity than GMW sequences. While high linear complexity is not a guarantee of cryptographic security, it does provide security against one of the few known generally effective cryptanalytic attacks on sequences, the Berlekamp-Massey algorithm. Moreover, cascaded GMW sequences are balanced and can be generated using standard shift register hardware, as will be discussed in one of the examples in Section A.

We will assume a basic understanding of finite fields and the trace function, since this material is very well explained in the excellent survey papers and books on the subject [7, 13, 15, 18]. Let q be a fixed power of 2 and let $GF(q)$ denote the Galois field with q elements. For any $n \geq 1$, we denote the *trace function* from $GF(q^n)$ to $GF(q)$ by $Tr_q^{q^n}$, defined by $Tr_q^{q^n}(x) = \sum_{i=0}^{n-1} x^{q^i}$. Recall that $Tr_q^{q^n}$ is a $GF(q)$-linear function, that every $GF(q)$-linear function f from $GF(q^n)$ to $GF(q)$ can be written in the form $f(x) = Tr_q^{q^n}(Ax)$, for some $A \in GF(q^n)$, and that, for any $m \geq 1$, $Tr_q^{q^{nm}}(x) = Tr_q^{q^n}(Tr_{q^n}^{q^{nm}}(x))$.

Let α be a primitive element of $GF(q^n)$, that is $GF(q^n)$ consists of zero and the powers of α. The infinite periodic sequence whose ith term is $Tr_q^{q^n}(\alpha^i)$ is known as an *m-sequence* over $GF(q)$ of span n [13]. (More generally, we can consider the sequence whose ith term is $Tr_q^{q^n}(A\alpha^i)$ for some fixed element A of $GF(q^n)$. This amounts to a cyclic shift of the first sequence, so we do not consider it to be a distinct sequence here.) It is well known that this sequence can be generated by a linear feedback shift register of length n over $GF(q)$. It has period $q^n - 1$, the maximum possible period for a sequence generated by a linear feedback shift register of length n over $GF(q)$. Moreover, every sequence with elements in $GF(q)$ which can be generated by a linear feedback shift register and has maximal period $q^n - 1$ is (a shift of) an m-sequence [13, pp. 394-410].

Suppose that $f : GF(q) \to GF(2)$ is a (possibly nonlinear) "feedforward" function. This, together with an m-sequence over $GF(q)$ of span n, gives rise to the *geometric sequence* [4] based on f and α,

$$\mathbf{S}_i = f(Tr_q^{q^n}(\alpha^i)).$$

A geometric sequence can be thought of as formed by applying a nonlinear feedforward function to the output of a maximal period linear feedback shift register. In particular, geometric sequences often are relatively simple to generate by hardware, as will be discussed in Section A. Geometric sequences are a very general class of pseudorandom binary sequences which includes the m-sequences (by taking $f(x) = Tr_2^q(x)$) and the GMW sequences (by taking $f(x) = Tr_2^q(x^k)$, where k is relatively prime to $q-1$) [17]. In other words, GMW(q^n, k, α) is the sequence

$$\mathbf{S}_i = Tr_2^q(Tr_q^{q^n}(\alpha^i)^k). \tag{1}$$

It is possible to generalize this construction if we have a tower of finite fields,

$$GF(2) \subset GF(q_1) \subset GF(q_2) \subset \ldots \subset GF(q_\ell),$$

by considering a nonlinear feedforward function $h : GF(q_\ell) \to GF(2)$ which is a composition of GMW-type functions, $h_i = Tr_{q_{i-1}}^{q_i}(x^{k_i})$ each of which drops down one step in the tower. We call such a composition $h = h_0 h_1 \cdots h_\ell$ a *cascade* and the resulting sequence $\mathbf{T}_i = h(\beta^i)$ is called a *cascaded GMW sequence*. An example of this construction, in the case $\ell = 3$, has been considered independently by Antweiler and Bömer [1].

In this paper we show that a cascaded GMW sequence of period $q_\ell - 1$ has the same autocorrelation function as that of an m-sequence of the same period, and that the linear complexity of the cascaded sequence can be made to grow exponentially with the number of steps in the cascade. Subsection A contains several examples of how the parameters for a cascaded GMW sequence can be chosen to make the linear complexity large. We also calculate the cross-correlation of a cascaded GMW sequence with an m-sequence of the same period, under the assumption that each of the steps $h_i : GF(q_i) \to GF(q_{i-1})$ is a quadratic function (that is, k_i has q_{i-1}-adic weight equal to 2, so h_i can be expressed as a polynomial of degree two in n_i variables over $GF(q_{i-1})$, where $q_i = q_{i-1}^{n_i}$). The techniques in this paper allow us to mix quadratic and linear functions, however the linear steps may all be removed without changing the sequence (as will be explained below). We find that the resulting cross-correlation function is three-valued and is low, just as in the case of Gold sequences and GMW sequences.

It is rather striking that such a computation can be made at all, because the cascaded feedforward function h is highly nonlinear and the complexity of the resulting sequence is enormous.

We conclude by combining the three properties – high linear complexity, minimal shifted autocorrelation, and low cross-correlation – in a single example. Specifically, for any integer ℓ, we exhibit a pair of binary sequences $\mathbf{S}$ and $\mathbf{T}$ with period $2^{5^\ell} - 1$, with linear complexities equal to $5^{2^{\ell-1}-1}$ and $5^{4^{\ell-1}-1}$, shifted autocorrelations equal to -1, and three valued cross-correlations, with values -1, $2^{(5^\ell+1)/2}-1$, and $-2^{(5^\ell+1)/2}-1$. Thus the cross-correlations are bounded by approximately the square root of two times the square root of the period, while the linear complexity is exponentially larger than the linear complexity of an m-sequence with the same period.

We would like to thank an anonymous referee for his or her careful reading of the first draft of this manuscript, and for her or his helpful comments and suggestions.

II Definitions

In this section we give the formal definition of the cascaded GMW sequences and establish notation which will be used in the rest of the paper.

Let $n_1, \ldots, n_\ell$ be positive integers, $q_0 = 2$, and $q_i = q_{i-1}^{n_i}$ (for $i = 1, \ldots, \ell,$). Let $k_1, \ldots, k_{\ell-1}$ be positive integers satisfying $k_i < q_i$ and $\gcd(k_i, q_i - 1) = 1$. Let α be a primitive element of $GF(q_\ell)$.

Definition 1 *The sequence* $\mathbf{T}$ *whose* j*th term is*

$$\mathbf{T}_j = Tr_{q_0}^{q_1}(Tr_{q_1}^{q_2}(\cdots Tr_{q_{\ell-1}}^{q_\ell}(\alpha^j)^{k_{\ell-1}}\cdots)^{k_1}) \tag{2}$$

is called a cascaded GMW sequence.

As we will see, choosing k_i to have q_{i-1}-adic weight two results in a large linear complexity. In some cases the linear complexity can be raised further by increasing the weight. It is unclear, however, how far we can go with this. It may be that there is no advantage in security to very large weights. There is an advantage in computational complexity, however, to keeping the weight small, since raising to elements to powers of two can be done quickly.

If $\ell = 1$, the sequence $\mathbf{T}$ is an m-sequence of period $q_1 - 1$. If $\ell = 2$, the sequence $\mathbf{T}$ is a GMW sequence. These are special cases of a more general construction: Suppose $f : GF(q) \to GF(2)$ is a (possibly nonlinear) "feedforward" function.

Definition 2 *The geometric sequence based on* f *and* α *is the sequence*

$$\mathbf{S}_i = f(Tr_q^{q^n}(\alpha^i)). \tag{3}$$

III Linear Complexity

In this section we study the linear complexity of a cascaded GMW sequence. In the case where each exponentiation is quadratic, the linear complexity is completely determined and is given in Theorem 1. We will make use of the following result of Brynielsson [3] (which follows from the work of Zierler and Mills on products of linear feedback shift register sequences [21]). This result can be applied to any geometric sequence based on a field of characteristic 2:

Proposition 1 *Let q be a power of 2. Suppose the sequence $\mathbf{S}$ is defined by equation (3) with the feedforward function $f : GF(q) \to GF(q)$ (as opposed to $f : GF(q) \to GF(2)$). Let f have a polynomial representation*

$$f = \sum_{i=0}^{q-1} A_i x^i$$

with coefficients A_i in $GF(q)$. Then the $GF(q)$-linear complexity of the sequence $\mathbf{S}$ is equal to $\sum_{A_i \neq 0} n^{||i||}$, where $||i||$ denotes the dyadic weight of the integer i (i.e. the number of ones in the base 2 representation of i).

In our case f maps into $GF(2)$ which can be considered a subset of $GF(q)$, so the proposition applies. The conclusion refers to the $GF(q)$-linear complexity. We are interested in the $GF(2)$ linear complexity. The following lemma shows that the two are equal.

Lemma 1 *Let $\mathbf{S}$ be a sequence of elements of $GF(2)$ (i.e., of zeros and ones). Since $GF(2) \subset GF(q)$ we may also consider $\mathbf{S}$ to be a sequence of elements in $GF(q)$. Then the $GF(q)$-linear complexity of $\mathbf{S}$ equals the $GF(2)$-linear complexity of $\mathbf{S}$.*

Proof: The $GF(q)$-linear complexity of $\mathbf{S}$ is the length, k, of the shortest linear recurrence $\mathbf{S}_{i+k} = \sum_{j=0}^{k-1} a_j \mathbf{S}_{i+j}$ satisfied by the sequence, with $a_j \in GF(q)$. If L is a $GF(2)$-linear function from $GF(q)$ to $GF(2)$ such that $L(1) = 1$ (such a function always exists), then applying L to the above linear recurrence gives $\mathbf{S}_{i+k} = \sum_{j=0}^{k-1} L(a_j)\mathbf{S}_{i+j}$. This is a $GF(2)$-linear recurrence of the same length. It follows that the $GF(2)$-linear complexity is less than or equal to the $GF(q)$-linear complexity.

Conversely, any linear recurrence with coefficients in $GF(2)$ can be thought of as a linear recurrence with coefficients in $GF(q)$, so the reverse inequality holds as well. $\square$

Thus to determine the linear complexity of a cascaded GMW sequence we must determine which monomials appear with nonzero coefficients in the corresponding feedforward function $g_{\ell-1} : GF(q_{\ell-1}) \to GF(2)$, given by $g_{\ell-1}(x) = Tr_{q_0}^{q_1}(Tr_{q_1}^{q_2}(\cdots Tr_{q_{\ell-2}}^{q_{\ell-1}}(x^{k_{\ell-1}})^{k_{\ell-2}} \cdots)^{k_1})$, and the dyadic weights of the corresponding exponents.

Lemma 2 *Let $n > 0$, q be a power of a prime, and A, B, C, D be nonzero sums of powers of q^n. If t is an integer such that $0 < t < n$ and $n/\gcd(n,t)$ is odd, then for all j, j' such that $0 \leq j, j' < n$,*

$$q^j A + q^{j+t} B = q^{j'} C + q^{j'+t} D$$

if and only if $A = C, B = D$ and $j = j'$.

Proof: Suppose $j < j'$. Then $q^j < q^{j'} < q^n$. Thus every term in $q^j A$ is distinct from every term in $q^{j'} C$, as well as from every term in $q^{j+t} B$. This implies that all terms of $q^j A$ appear in $q^{j'+t} D$. Since A is nonzero, there is at least one such term. It follows that $j \equiv j' + t \bmod n$, hence, $j = j' + t - n$. Similarly, terms of $q^{j'} C$ appear in $q^{j+t} B$ so $j' \equiv j + t \bmod n$.

It follows that $2t \equiv 0 \bmod n$, so $t \equiv n/2 \bmod n$. Thus $n/\gcd(n,t) = 2$, contradicting the hypothesis that $n/\gcd(n,t)$ is odd. Therefore, $j = j'$ and $A = C, B = D$. $\square$

Theorem 1 *For each $i, 1 \le i \le \ell - 1$, let $k_i = q_{i-1}^{s_i} + q_{i-1}^{t_i}$ (i.e., k_i has q_{i-1}-adic weight two), with $0 \le s_i < t_i < n_i$ and $n_i/\gcd(n_i, t_i - s_i)$ odd. Then the linear complexity of the cascaded GMW sequence $\mathbf{T}$ is given by*

$$n_1 n_2^2 n_3^4 \cdots n_\ell^{2^{\ell-1}},$$

where $q_i = q_{i-1}^{n_i}, i = 1, \cdots, \ell$.

Proof: To prove the theorem, we need to show

1. the number of monomials with nonzero coefficients appearing in $g_{\ell-1}(x)$ is given by

$$n_1 n_2^2 n_3^4 \cdots n_{\ell-1}^{2^{\ell-2}},$$

and

2. the dyadic weight of each of the exponents appearing in the above monomials is $2^{\ell-1}$.

Then, by Brynielsson's result (Proposition 1) we see that the linear complexity of the cascaded GMW sequence $\mathbf{T}$ is given by

$$\begin{aligned} \sum_{A_i \neq 0} n_\ell^{||i||} &= (\text{number of nonzero } A_i\text{'s}) n_\ell^{2^{\ell-1}} \\ &= (n_1 n_2^2 n_3^{2^2} \cdots n_{\ell-1}^{2^{\ell-2}})(n_\ell^{2^{\ell-1}}). \end{aligned}$$

To show (1), let

$$\pi_{\ell-i-1}(x) = Tr_{q_i}^{q_{i+1}}(Tr_{q_{i+1}}^{q_{i+2}}(\cdots Tr_{q_{\ell-2}}^{q_{\ell-1}}(x^{k_{\ell-1}})^{k_{\ell-2}} \cdots)^{k_{i+1}}),$$

and $\pi_0(x) = x$, then $\pi_{\ell-i} = Tr_{q_{i-1}}^{q_i}(\pi_{\ell-i-1}^{k_i})$ and $\pi_{\ell-1}(x) = g_{\ell-1}(x)$. The polynomial representation of $g_{\ell-1}$ is derived by alternately raising $\pi_{\ell-1-i}$ to a k_ith power and applying a trace function. Note that raising to a power of 2 at any stage has no effect on the number of monomials or dyadic weights of exponents, so we may assume that $s_i = 0$ for each i. Thus $n_i/\gcd(n_i, t_i)$ is assumed odd.

We will show by induction that for each $i = \ell-1, \cdots, 1$, $\pi_{\ell-i}$ consists of $n_i n_{i+1}^2 n_{i+2}^4 \cdots n_{\ell-1}^{2^{\ell-i-1}}$ monomials, each of whose exponents is a sum of powers of q_{i-1}. As base case, we have

$\pi_1 = Tr^{q_{\ell-1}}_{q_{\ell-2}}(x^{k_{\ell-1}}) = \sum_{j=0}^{n_{\ell-1}-1} x^{k_{\ell-1}q^j_{\ell-2}}$. Assume inductively that, $\pi_{\ell-i-1}(x) = \sum_A x^A$, where each A is a nonzero sum of powers of q_i. Thus,

$$\begin{aligned}
\pi_{\ell-i}(x) &= Tr^{q_i}_{q_{i-1}}((\pi_{\ell-i-1}(x))^{k_i}) && (4)\\
&= Tr^{q_i}_{q_{i-1}}((\sum_A x^A)(\sum_B x^B)^{q^{t_i}_{i-1}}) && (5)\\
&= Tr^{q_i}_{q_{i-1}}(\sum_{A,B} x^{A+Bq^{t_i}_{i-1}}) && (6)\\
&= \sum_{j=0}^{n_{i-1}} \sum_{A,B} x^{(A+Bq^{t_i}_{i-1})q^j_{i-1}}. && (7)
\end{aligned}$$

By Lemma 2 all monomials in (7) are distinct, thus the number of monomials in $\pi_{\ell-i}$ equals n_i times the square of the number of monomials in $\pi_{\ell-i-1}$. Each $(A + Bq^{t_i}_{i-1})q^j_{i-1}$ is nonzero, so the induction goes through. For $i = 1$, we have that the number of monomials in $\pi_{\ell-1}$ is $n_1 n_2^2 n_3^{2^2} \cdots n_{\ell-1}^{2^{\ell-2}}$.

To get the last term $n_\ell^{2^{\ell-1}}$ in the expression of the linear complexity, we show by induction that all degrees of monomials of $\pi_{\ell-i}$ have the same dyadic weight $2^{\ell-i}$.

The degree of the monomial in $\pi_0(x)$ has dyadic weight 1. As we saw, the degrees of the monomials in $\pi_{\ell-i}(x)$ are of the form $C = q^j_{i-1}(A + Bq^{t_i}_{i-1})$, where A and B are degrees of monomials in $\pi_{\ell-i-1}$, hence sums of $2^{\ell-i-1}$ powers of q_i. Since $t_i \neq 0$, no term of A can appear in $Bq^{t_i}_{i-1}$, hence the C is a sum of $2^{\ell-i}$ powers of q_{i-1}, i.e., has dyadic weight $2^{\ell-i}$. □

By a similar proof we can show the following.

Theorem 2 *For each $i, 1 \le i \le \ell-1$, let $k_i = (q^{s_i}_{i-1}+q^{t_i}_{i-1})(q^{u_i}_{i-1}+q^{v_i}_{i-1})$ with $0 \le s_i < t_i < n_i$, $0 \le u_i < v_i < n_i$, $n_i/\gcd(n_i, t_i - s_i)$ odd, $n_i/\gcd(n_i, v_i - u_i)$ odd, $t_i - s_i < v_i - u_i$, and $n_i \neq 2(v_i - u_i)$. Then the linear complexity of the cascaded GMW sequence $\mathbf{T}$ is given by*

$$n_1 n_2^4 n_3^{4^2} \cdots n_\ell^{4^{\ell-1}},$$

where $q_i = q^{n_i}_{i-1}, i = 1, \cdots, \ell$.

More generally, we can mix the hypotheses of these theorems for different i, with a contribution of 2 to the exponents for each index i such that the hypotheses of Theorem 1 hold, and a contribution of 4 to the exponents for each index i such that the hypotheses of Theorem 2 hold.

A Examples

Example 1. If the exponents k_i or the extensions n_i are chosen carefully, the linear complexity of a cascaded GMW sequence can be made exponentially large. Suppose first that we restrict attention to quadratic exponents, as in Theorem 1 (in part because these exponents are simpler to implement, and in part because the analysis is tractable in this case). Consider a tower of fields with $n_1 = n_2 = \cdots = n_\ell = 3$. The highest linear complexity we can achieve for a GMW sequence with these restrictions is $3^{\ell+1}$, which occurs for the sequence with feedforward function

$$Tr_2^{2^{3^{\ell-1}}}(x^3).$$

In contrast, if we take a cascaded GMW sequence over this tower of fields, with $k_i = 1 + q_{i-1}$, then by Theorem 1, the linear complexity is

$$3 \cdot 3^2 \cdot 3^4 \cdots 3^{2^{\ell-1}} = 3^{2^\ell - 1},$$

which is exponentially larger. (Note that we are forced to take $n_i = 3$ rather than $n_i = 2$ in order to obtain a quadratic exponent which satisfies the hypotheses of the theorem.) The amount of hardware required in this case is $\mathcal{O}(\ell 3^{2\ell})$. Thus the linear complexity is exponentially larger than the hardware requirements.

We see the contrast in complexity between the GMW and cascaded GMW sequences more strikingly by taking $\ell = 3$. For the GMW sequence we have the feedforward functions

$$g(x) = Tr_2^{2^9}(x^3) = x^3 + x^6 + x^{12} + x^{24} + x^{48} + x^{96} + x^{192} + x^{384} + x^{257}$$

while the feedforward function for the cascaded GMW sequence is

$$\begin{aligned} g_2(x) &= Tr_2^{2^3}(Tr_{2^3}^{2^9}(x^9)^3) \\ &= x^{27} + x^{45} + x^{54} + x^{83} + x^{90} + x^{101} + x^{108} + x^{139} + x^{153} + x^{166} + x^{180} + x^{195} \\ &\quad + x^{202} + x^{209} + x^{216} + x^{269} + x^{278} + x^{297} + x^{306} + x^{325} + x^{332} + x^{353} + x^{360} \\ &\quad + x^{390} + x^{404} + x^{418} + x^{432}. \end{aligned}$$

Clearly, the latter is more complex, both in the number of terms and in the dyadic weights of the exponents of the terms (2 for the former, 4 for the later). Yet the computation of g_2 is hardly more complex than that of g. This cascaded GMW sequence can be generated as follows. We first establish a primitive normal basis for $GF(2^9)$. That is, we select a primitive element α such that $\{\alpha, \alpha^2, \alpha^4, \cdots, \alpha^{2^8}\}$ are linearly independent, hence forms a so-called normal basis, over $GF(2)$. Such a basis is known to always exist [12]. In this representation, the element $\sum_{i=0}^{8} a_i \alpha^{2^i}$ corresponds to the binary sequence $(a_0, \cdots, a_8)$. Moreover, addition corresponds to bitwise exclusive or, and squaring corresponds to cyclic shift to the right one position. Multiplication can be implemented by a log-antilog table. This consists of a pair of arrays, *log* and *alog*, of size 2^9. These tables are defined as follows: $x = \alpha^i$ if and only if $log[x] = i$ and $alog[i] = x$ for $x \neq 0$. For nonzero x and y in $GF(2^9)$, we can then compute the product of x and y as $alog[log[x] + log[y] \bmod (2^9 - 1)]$. The cases where x or

y is zero can be treated specially. Once such tables are established, addition, exponentiation to a power of 2, and multiplication are very fast (the addition of 9 bit integers used in multiplication is the slowest operation in general).

To represent elements of $GF(2^3)$, observe the following. First, $\beta = \alpha + \alpha^8 + \alpha^{64}$ is a primitive element of $GF(2^3)$. Second, an element of $GF(2^9)$ is in $GF(2^3)$ if and only if bits 0, 3, and 6 are equal, bits 1, 4, and 7 are equal, and bits 2, 5, and 8 are equal (this amounts to saying $x^8 = x$, which characterizes $GF(2^3)$). Thus the mapping from the representation of an element of $GF(2^3)$ as an element of $GF(2^9)$ using the normal basis generated by α to the representation using the normal basis generated by β consists of projection on the first three coordinates. It follows that to compute a $Tr_{2^3}^{2^9}(x)$, we need only compute the first three bits of $x + x^8 + x^{64}$. These are the mod two sum of bits 0, 3, and 6, of bits 1, 4, and 7, and of bits 2, 5, and 8, respectively. This trace operation can therefore be built in very simple hardware (with six exclusive or gates). Calculation of $y = Tr_{2^3}^{2^9}(x^9)$ requires an additional exponentiation by a power of 2 to compute x^8, and a multiplication by x to compute x^9. Calculation of $Tr_2^{2^3}(y^3)$ then requires one squaring and one multiplication in $GF(2^3)$, and two exclusive or gates.

Finally, we must generate the original m-sequence over $GF(2^9)$. As mentioned in the introduction, this can be done using a linear feedback shift register of length three over $GF(2^9)$. If the elements of the register are x_1, x_2, and x_3, then the register operates by outputing x_1 and replacing x_1 by x_2, x_2 by x_3, and x_3 by $a_1x_1 + a_2x_2 + a_3x_3$ at each iteration, where a_1, a_2, and a_3 are elements of $GF(2^9)$. This involves three multiplications and two additions in $GF(2^9)$, in addition to simple hardware to perform the shifting. This can be simplified if we can find such a register where one or more of the coefficients is zero or one.

Similar considerations can be applied in general to the generation of cascaded GMW sequences. If the field $GF(q_\ell)$ is large, however, it may be impractical to store the *log* and *alog* tables. It may also be difficult to find a primitive normal basis. We can, however, use a (non-primitive) normal basis and implement multiplication by circuitry. This can always be done with approximately $4n \log n$ gates, where $q_\ell = 2^n$. Such a basis can always be found [13]. In this case minimizing the number of multiplications is more critical. To improve efficiency, for some fields we can use the fast multiplication techniques due to Mullin, Onyszchuk, and Vanstone [16].

Example 2. Suppose we leave the exponents unrestricted. Again consider a tower of fields with $n_1 = n_2 = \cdots = n_\ell = 3$, so $q_i = 2^{3^i}$. Thus the periods of the sequences generated is at most $2^{3^\ell} - 1$. The largest possible linear complexity for a GMW sequence with this period is $3^{3^{\ell-1}+\ell-2}$. This occurs for the sequence with feedforward function

$$Tr_2^{2^{3^{\ell-1}}}(x^{2^{(3^{\ell-1}-1)}-1}).$$

Now consider a cascaded GMW sequence with $k_i = 1 + q_{i-1}$ for $i = 1, \cdots, \ell - 2$, and $k_{\ell-1}$ the solution to the congruence

$$k_{\ell-1} \prod_{i=0}^{\ell-1} (1 + q_i) \equiv 2^{(3^{\ell-1}-1)} - 1 \bmod 2^{3^{\ell-1}} - 1.$$

It follows that the product of the k_i's has $q_{\ell-1}$-adic weight $3^{\ell-1} - 1$, which is the maximum weight exponent that can contribute to the final feedforward function. Further, we note that $\gcd(2^{3^i} + 1, 2^{3^{\ell-1}} - 1) = 1$, so the above congruence can be solved uniquely.

Let

$$g(x) = Tr_2^{2^3}(Tr(_{2^3}^{2^9}(\cdots Tr_{2^{3^{\ell-2}}}^{2^{3^{\ell-1}}}(x)^{k_{\ell-2}} \cdots)^{k_2})^{k_1}).$$

The feedforward function defined by these k_i is $f(x) = g(x^{k_{\ell-1}})$. The argument in the proof of Theorem 1 shows that there is no cancellation of terms in $g(x)$ as we perform the compositions of traces and exponentiations. Moreover, $k_{\ell-1}$ is relatively prime to $q_{\ell-1} - 1$, so raising to the $k_{\ell-1}$ power is a permutation of $GF(q_{\ell-1})$. It follows that composing $g(x)$ with $x^{k_{\ell-1}}$ simply changes the exponents, that is, does not introduce any cancellation. Thus all terms in $f(x)$ remain distinct.

We also know that f maps $GF(q_{\ell-1})$ to $GF(2)$. It follows from Galois theory that if f has a term x^u, then it also has every term x^v where $v \equiv 2^j u \pmod{q_{\ell-1} - 1}$ for some j. By construction, f has the term x^u, $u = 2^{(3^{\ell-1}-1)} - 1$, and hence has all $3^{\ell-1}$ terms whose exponents have dyadic weight $3^{\ell-1} - 1$. These terms contribute $3^{3^{\ell-1}+\ell-2}$ to the linear complexity, and this is precisely the linear complexity of the GMW sequence we discussed above. Thus the linear complexity of this cascaded sequence exceeds the linear complexity of the GMW sequence by $\sum 3^{||j||}$, where the sum ranges over the remaining $3^{2^{\ell-1}-1} - 3^{\ell-1}$ terms, hence by at least $3^{2^{\ell-1}} - 3^{\ell-1}$.

IV AUTOCORRELATION

Recall that the (periodic) cross-correlation function of two binary sequences $\mathbf{S}$ and $\mathbf{T}$ of period $q^n - 1$ is defined as

$$\Theta_{\mathbf{S},\mathbf{T}}(\tau) = \sum_{r=1}^{q^n-1} (-1)^{\mathbf{S}_{r+\tau}} (-1)^{\mathbf{T}_r}.$$

and the autocorrelation function of the sequence $\mathbf{S}$ is defined as $A_{\mathbf{S}}(\tau) = \Theta_{\mathbf{S},\mathbf{S}}(\tau)$. For any geometric sequence, $\mathbf{S}_i = f(Tr_q^{q^n}(\beta^i))$ the authors previously showed [10] that the autocorrelation function is given by

$$A_{\mathbf{S}}(\tau) = q^{n-1} c_f(t) - 1 \qquad \text{if } \tau = t\nu, \tag{8}$$
$$A_{\mathbf{S}}(\tau) = q^{n-2} I(f)^2 - 1 \qquad \text{otherwise}. \tag{9}$$

where the imbalance $I(f) = \sum_{u \in GF(q)} (-1)^{f(u)}$, $\nu = (q^n - 1)/(q-1)$ and where $c_f(t)$ is the autocorrelation function of f,

$$c_f(t) = \sum_{u \in GF(q)} (-1)^{f(u)} (-1)^{f(\beta^t u)}.$$

Here, $\beta = \alpha^\nu$ is the corresponding primitive element of $GF(q)$.

In the case of a GMW sequence, $f(x) = Tr_2^q(x^k)$ and $I(f) = 0$. Since $\gcd(k, q-1) = 1$, $x \to x^k$ is a permutation on $GF(q)$. Hence

$$\begin{aligned} c_f(t) &= \sum_{u \in GF(q)} (-1)^{Tr_2^q(u^k + (\beta^t u)^k)} \\ &= \sum_{v \in GF(q)} (-1)^{Tr_2^q(v + \beta^{tk} v)} \\ &= \sum_{v \in GF(q)} (-1)^{Tr_2^q((1+\beta^{tk})v)} \\ &= \begin{cases} q & \text{if } t = 0 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

It follows that (c.f. [17])

$$A_{\mathbf{S}}(\tau) = \begin{cases} q^n - 1 & \text{if } \tau = 0 \\ -1 & \text{otherwise.} \end{cases}$$

We can use this result as the base case of an induction to calculate the autocorrelation function of a cascaded GMW sequence by applying equation (8). The imbalance is again zero, so we obtain:

Proposition 2 *The autocorrelation function of the cascaded GMW sequence (1) is given by*

$$A_{\mathbf{S}}(\tau) = \begin{cases} q_\ell - 1 & \text{if } \tau = 0 \\ -1 & \text{otherwise.} \end{cases}$$

V Cross-correlation

The authors previously calculated the periodic cross-correlation function for "quadratically decimated" geometric sequences [10]. We recall the main results here, so they may be applied to the case of cascaded GMW sequences.

Suppose that α and β are primitive elements of $GF(q^n)$, and that $\beta = \alpha^k$. If we are given two geometric sequences,

$$\mathbf{S}_r = f(Tr_q^{q^n}(\alpha^r)) \text{ and } \mathbf{T}_r = g(Tr_q^{q^n}(\beta^r)) = g(Tr_q^{q^n}(\alpha^{kr})) \tag{10}$$

we say that the sequence $\mathbf{T}$ is related to the sequence $\mathbf{S}$ by a decimation k (and a change of feedforward function). If $k = 2^e$ we say the decimation is *linear*. If $k = q^i + q^j$ we say the decimation is *quadratic*. The periodic cross-correlation of $\mathbf{S}$ and $\mathbf{T}$ is given by the following results, and it involves the *short cross-correlation* function,

$$\Delta_a^e(f,g) = \sum_{u \in GF(q)} F(au)G(u^{2^e})$$

where $F(u) = (-1)^{f(u)}$, $G(u) = (-1)^{g(u)}$ and $a \in GF(q)$.

Theorem 3 *[10] Let $\mathbf{S}$ and $\mathbf{T}$ denote the geometric sequences defined in equation (10). Suppose $\mathbf{T}$ is quadratically related to $\mathbf{S}$ with decimation $k = q^i + q^j$. Let $m = n - \gcd(n, j-i)$. Then m is even and the values for the cross-correlation function $\Theta_{\mathbf{S},\mathbf{T}}(\tau)$ are:*

1. $\Theta_{\mathbf{S},\mathbf{T}}(\tau) = q^{n-2}I(f)I(g) - F(0)G(0)$ *which occurs for* $q^n - q^{m+1} + q^m - 1$ *values of* τ.
2. $\Theta_{\mathbf{S},\mathbf{T}}(\tau) = (q^{n-2} - q^{n-m/2-2})I(f)I(g) + q^{n-m/2-1}\Delta_a^1(f,g) - F(0)G(0)$ *which occurs for* $(q^m + q^{m/2})/2$ *values of* τ *for each nonzero* $a \in GF(q)$.
3. $\Theta_{\mathbf{S},\mathbf{T}}(\tau) = (q^{n-2} + q^{n-m/2-2})I(f)I(g) - q^{n-m/2-1}\Delta_a^1(f,g) - F(0)G(0)$ *which occurs for* $(q^m - q^{m/2})/2$ *values of* τ *for each nonzero* $a \in GF(q)$.

Theorem 4 *[10] Let $\mathbf{S}$ and $\mathbf{T}$ denote the geometric sequences defined in equation (10). Suppose that they are related by a linear decimation $k = 2^e$. Then the values for the cross-correlation are:*

1. $\Theta_{\mathbf{S},\mathbf{T}}(\tau) = q^{n-2}I(f)I(g) - F(0)G(0)$ *which occurs for* $q^n - q$ *values of* τ, *whenever* $\alpha^\tau \notin GF(q)$
2. $\Theta_{\mathbf{S},\mathbf{T}}(\tau) = q^{n-1}\Delta_a^e(f,g) - F(0)G(0)$ *which occurs once for each nonzero value of* $a = \alpha^{-\tau} \in GF(q)$

In order to apply Theorem 3 to cascaded GMW sequences, we consider the case in which, for each $i \geq 1$, k_i has q_{i-1}-adic weight two, and we write $k_i = q_{i-1}^{s_i} + q_{i-1}^{t_i}$, where $0 \leq s_i \leq t_i < n_i$, $s_1 < t_1$.

Note that if any $k_i = 2^a$ has dyadic weight 1, then this step may be eliminated from the cascade because $Tr_{q_{i-1}}^{q_i}(x^2) = (Tr_{q_{i-1}}^{q_i}(x))^2$ so the power of 2 may be inductively moved down the cascade until it arrives at the bottom, where it acts as the identity. Thus the case in which there is a mix of weight two and weight one exponents reduces to a smaller case in which all the exponents have weight two.

As with GMW sequences, the cross-correlation function of an m-sequence and a cascaded GMW sequence is three valued. We let ℓ, $\{n_i\}$, $\{q_i\}$, $\{k_i\}$ determine a cascaded GMW sequence $\mathbf{T}$ with primitive element β, as in equation 2.

Theorem 5 *Let* $\mathbf{T}$ *be a cascaded GMW sequence (as in Equation (2)), and let* $\mathbf{S}$ *be the m-sequence* $Tr_2^{q_\ell}(\alpha^{i-1})$*, where* α *is a primitive element of* $GF(q_\ell)$ *over* $GF(2)$*. Let* $\beta = \alpha^{k_\ell}$*. Assume that, for each* $i \geq 1$*,* $k_i = q_{i-1}^{s_i} + q_{i-i}^{t_i}$ *for* $0 \leq s_i \leq t_i < n_i$*,* $s_1 < t_1$*, i.e.,* k_i *has* q_{i-1}*-adic weight 2. Assume moreover that, for each* i*,* $n_i/\gcd(n_i, t_i - s_i)$ *is odd (i.e.,* $\gcd(k_i, q_i - 1) = 1$*). Let* $m_i = n_i - \gcd(n_i, t_i - s_i)$*,* $d_i = n_1 n_2 \cdots n_i$ *(*$d_0 = 1$*), and* $e_i = \sum_{j=1}^{i} d_{j-1} m_j$*. Then the cross-correlation function of* $\mathbf{S}$ *and* $\mathbf{T}$ *is given by*

$$\Theta_{\mathbf{S},\mathbf{T}}(\tau) = \begin{cases} -1 & \textit{occurring} \quad 2^{d_\ell} - 2^{e_\ell} - 1 & \textit{times} \\ 2^{d_\ell - e_\ell/2} - 1 & \textit{occurring} \quad (2^{e_\ell} + 2^{e_\ell/2})/2 & \textit{times} \\ -2^{d_\ell - e_\ell/2} - 1 & \textit{occurring} \quad (2^{e_\ell} - 2^{e_\ell/2})/2 & \textit{times.} \end{cases}$$

Proof: We proceed by induction on ℓ. The base case of the induction ($\ell = 1$) is given by Theorem 1 with $f = g = id : GF(2) \to GF(2)$, $q = 2, n = n_1, m = m_1$. Note also that $\Delta_1^1(f, g) = 2$. To see the general case, we define for each i

$$f_i(x) = Tr_2^{q_i}(x)$$

$$g_i(x) = Tr_2^{q_1}(Tr_{q_1}^{q_2}(\cdots Tr_{q_{i-1}}^{q_i}(x^{k_i})^{k_{i-1}} \cdots)^{k_1}).$$

Thus $\mathbf{S}$ and $\mathbf{T}$ are the sequences whose ith terms are $f_\ell(\alpha^{i-1})$ and $g_\ell(\alpha^{i-1})$ respectively. In particular, they are geometric sequences with feedforward functions $f_{\ell-1}$ and $g_{\ell-1}$, respectively. The function $f_{\ell-1}$ has zero imbalance, while $f_{\ell-1}(0) = g_{\ell-1}(0) = 0$. Applying Theorem 3 we obtain

1. -1 occurs $q_{\ell-1}^{n_\ell} - q_{\ell-1}^{m_\ell+1} + q_{\ell-1}^{m_\ell} - 1$ times,
2. $q_{\ell-1}^{n_\ell - m_\ell/2-1} \Delta_a^1(f_{\ell-1}, g_{\ell-1}) - 1$ occurs $(q_{\ell-1}^{m_\ell} + q_{\ell-1}^{m_\ell/2})/2$ times, and
3. $-q_{\ell-1}^{n_\ell - m_\ell/2-1} \Delta_a^1(f_{\ell-1}, g_{\ell-1}) - 1$ occurs $(q_{\ell-1}^{m_\ell} - q_{\ell-1}^{m_\ell/2})/2$ times.

Here

$$\begin{aligned} \Delta_a^1(f_{\ell-1}, g_{\ell-1}) &= \sum_{u \in GF(q_{\ell-1})} (-1)^{f_{\ell-1}(u) + g_{\ell-1}(u^2/a^2)} \\ &= \sum_{u \in GF(q_{\ell-1})} (-1)^{f_{\ell-1}(au) + g_{\ell-1}(u^2)} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{u \in GF(q_{\ell-1})} (-1)^{f_{\ell-1}(au)+g_{\ell-1}(u)^2} \\
&= \sum_{u \in GF(q_{\ell-1})} (-1)^{f_{\ell-1}(au)+g_{\ell-1}(u)} \\
&= \Theta_{\mathbf{S}',\mathbf{T}'}(\sigma) + 1,
\end{aligned}
$$

where $\mathbf{S}'$ and $\mathbf{T}'$ are the sequences whose jth terms are $f_{\ell-1}(\gamma^{j-1})$ and $g_{\ell-1}(\gamma^{(j-1)})$, respectively, γ any primitive element of $GF(q_{\ell-1})$, and $\gamma^\sigma = a$ (this follows from the fact that $Tr^q_{q'}(u^2) = Tr^q_{q'}(u)^2$ in characteristic two, and $g_{\ell-1}(u)$ is in $GF(2)$, hence equal to its square). $\mathbf{S}'$ is a m-sequence, and $\mathbf{T}'$ a cascaded GMW sequence, with one less level than $\mathbf{T}$, and $\mathbf{S}'$ and $\mathbf{T}'$ satisfy the hypotheses of the theorem. By induction, the theorem holds for $\mathbf{S}'$ and $\mathbf{T}'$. This in turn implies the theorem for $\mathbf{S}$ and $\mathbf{T}$. □

More generally, suppose $\mathbf{S}$ and $\mathbf{T}$ are both cascaded GMW sequences over the same tower of fields, with exponents $\{k_i\}$ and $\{k'_i\}$, respectively. If for each i, either $k'_i/k_i = q_{i-1}^{s_i} + q_{i-1}^{t_i}$ or $k_i/k'_i = q_{i-1}^{s_i} + q_{i-1}^{t_i}$, and the hypotheses of Theorem 5 hold with respect to $\{s_i\}$ and $\{t_i\}$, then the conclusions of the theorem hold for the cross-correlation $\Theta_{\mathbf{S},\mathbf{T}}$ of $\mathbf{S}$ and $\mathbf{T}$.

VI Summary and Conclusions

We have defined a new class of periodic binary sequences, the cascaded GMW sequences, based on a tower of finite fields. We have shown that these sequences have minimal shifted autocorrelations (Section IV), can have large linear complexities (larger than those of GMW sequences – Section III), and can have low three valued cross-correlations (Section V).

Example 3. We conclude with a single example of a pair of sequences with all these properties. Let each $n_i = 5$, to form a tower of finite fields of characteristic two, as in the definition of cascaded GMW sequences, and let $k_i = 1 + q_i$, and $k'_i = (1+q_i)(1+q_i^2)$ be the exponents used to form a pair of cascaded GMW sequences over this tower. The hypotheses of Theorem 1 hold for $\mathbf{S}$, the hypotheses of Theorem 2 hold for $\mathbf{T}$, and the hypotheses of Theorem 5 hold. Thus we have a pair of sequences of period $2^{5^\ell} - 1$, with linear complexities $5^{2^\ell - 1}$ and $5^{4^\ell - 1}$, shifted autocorrelations equal to -1, and three valued cross-correlation function, with values -1, $2^{(5^\ell+1)/2} - 1$, and $-2^{(5^\ell+1)/2} - 1$. Note that we can increase the minimum of the linear complexities of these two sequences without changing the correlation properties by interchanging the values of k_i and k'_i for every other value of i. The details are left to the reader.

For $\ell = 3$, by interchanging alternate values of k_i and k'_i, we obtain a pair of sequences of period $2^{125} - 1$ with linear complexities 5^{11} and 5^{13}, shifted autocorrelations equal to -1, and three valued cross-correlation with values -1, $2^{63} - 1$ and $-2^{63} - 1$. These are high linear complexities, optimal autocorrelations, and near optimal cross-correlations for sequences of this period. Moreover, these sequences can be generated by relatively simple shift register hardware.

A number of questions remain. We would like to find larger families of sequences that have high linear complexity and low correlation values. Unfortunately, if $\mathbf{S}$, $\mathbf{T}$, and $\mathbf{U}$ are three cascaded GMW sequences, where $\mathbf{S}$ and $\mathbf{T}$ have quadratically related exponents and $\mathbf{T}$ and $\mathbf{U}$ have quadratically related exponents, it may not be the case that $\mathbf{S}$ and $\mathbf{U}$ have quadratically related exponents. It is thus desirable to extend the results of Theorem 5 (or, more generally, of Theorem 3) to non-quadratic exponents. This appears to require quite deep results from algebraic geometry that are as yet unavailable.

Another concern is that there is still a large gap between the linear complexities and the periods of cascaded GMW sequences. Brynielsson's result implies that the linear complexity is at most $(n+1)^e - 2^e$, where $q = 2^e$. We have not quite achieved this, and it is concievable that we can do better, say by tightening the estimate in the final example of Section III (here the bound given by Brynielsson's result is $4^{3^{\ell-1}} - 2^{3^{\ell-1}}$, while we find a linear complexity of at least $3^{3^{\ell-1}+\ell-2} + 3^{2^{\ell-1}-1} - 3^{\ell-1}$). Nonetheless, these results are a positive step toward producing cryptographically secure pseudo-random sequences with good correlation properties.

References

[1] M. Antweiler and L. Bömer, "Complex Sequences over $GF(p^m)$ with a Two-Level Autocorrelation Function and a Large Linear Span," manuscript, 1990.

[2] O. Rothaus, "On bent functions," *Journal of Combinatorial Theory,* Series A, Vol. 20, pp. 300-305, 1976.

[3] L. Brynielsson, "On the linear complexity of combined shift registers," *Proceedings of Eurocrypt 1985,* published in *Advances in Cryptology – Eurocrypt '85, Lecture Notes in Computer Science, Vol. 219,* pp. 156-166, Springer-Verlag, Berlin, 1985.

[4] A. H. Chan and R. Games, "On the linear span of binary sequences from finite geometries, q odd," *Proceedings of Crypto 1986,* published in *Advances in Cryptology – Crypto '86, Lecture Notes in Computer Science, Vol. 263,* pp. 405-417, Springer-Verlag, Berlin, 1987.

[5] R. Games, "Crosscorrelation of m-sequences and GMW- sequences with the same primitive polynomial," *Discrete Applied Mathematics,* Vol. 12, pp. 139-146, 1985.

[6] R. Games "The geometry of m-sequences: three-valued cross-correlations and quadrics in finite projective geometry," *SIAM J. Alg. Disc. Methods,* Vol. 7, pp. 43-52, 1986.

[7] S. Golomb, *Shift Register Sequences,* Aegean Park Press, Laguna Hills, CA, 1982.

[8] B. Gordon, W. H. Mills, and L. R. Welch, "Some new difference sets," *Canad. J. Math.,* Vol. 14, pp. 614-625, 1962.

[9] T. Helleseth, "Some results about the cross-correlation function between two maximal linear sequences," *Discrete Math,* Vol. 16, pp. 209-232, 1976.

[10] A. Klapper, A. H. Chan, and M. Goresky, "Cross-correlations of linearly and quadratically related geometric sequences and GMW sequences," to appear, *Discrete Applied Mathematics.*

[11] V. Kumar and O. Moreno, "Polyphase Sequences with Periodic Correlation Properties Better than Binary Sequences," manuscript, 1990.

[12] H. Lenstra, Jr. and R. Schoof, "Primitive normal bases for finite fields," *Mathematics of Computation,* Vol. 48, No. 177, pp. 217-231, 1987.

[13] R. Lidl and H. Niederreiter *Finite Fields* in *Encyclopedia of Mathematics, Volume 20,* Cambridge University Press, Cambridge, 1983.

[14] J. L. Massey, "Shift register synthesis and BCH decoding," *IEEE Trans. Info. Thy.,* Vol. IT-15, pp. 122-127, 1969.

[15] R. McEliece, *Finite Fields for Computer Scientists and Engineers,* Kluwer Academic Publishers, Boston, 1987.

[16] R. Mullin, I. Onyszchuk, and S. Vanstone, "Optimal normal bases in $GF(p^n)$," *Discrete Applied Mathematics,* Vol. 22, pp. 149-161, 1989.

[17] R. Scholtz and L. Welch, "GMW sequences," *IEEE Trans. Info. Thy.,* Vol. IT-30, No. 3, pp. 548-553, 1984.

[18] M. Simon, J. Omura, R. Scholtz, and B. Levitt, *Spread-Spectrum Communications, Volume 1*, Computer Science Press, 1985.

[19] D. Sarwate and M. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *IEEE Proceedings,* Vol. 68, pp. 593-619, 1980.

[20] L. R. Welch, "Lower bounds on the maximum correlation of signals," *IEEE Trans. Inform. Theory,* Vol. IT-20, pp. 397-399, 1974.

[21] N. Zierler and W. H. Mills, "Products of linear recurring sequences," *J. of Algebra*, Vol. 27, pp. 147-157, 1973.