Large Families of Sequences with Near Optimal Correlations and Large Linear Span

Andrew Klapper*
Department of Computer Science
763H Anderson Hall
University of Kentucky
Lexington, KY 40506-0046
E-mail: klapper@cs.engr.uky.edu

Abstract

In order to build spread spectrum communication systems based on the CDMA paradigm, it is necessary to have large families of binary sequences with low pairwise correlation values. For these systems to have resistance to certain cryptanalytic attacks and resistance to jamming, the sequences must have large linear span. In this paper we describe certain families of sequences that have these desirable properties. The sequences are based on families of quadratic forms over finite fields.

1 Introduction

In recent years there has been hope that Code Division Multiple Access (CDMA) systems would provide the capacity and reliability necessary to make spread spectrum communications viable [13]. The volume of communication traffic has been steadily increasing, and will continue to do so. In order for these systems to work, however, it is essential to find large families of easily generated binary sequences with low correlation function values. The smaller the pairwise cross-correlations and the larger the family, the higher the capacity of the system. For these systems to resist jamming and certain cryptanalytic attacks, the sequences must have large linear spans.

*Project sponsored by the National Security Agency under Grant Number MDA904-91-H-0012 and the National Science Foundation under Grant Number NCR-9400762. The United States Government is authorized to reproduce and distribute reprints notwithstanding any copyright notation hereon.

		Size of	Maximum	Maximum	Range of
Family	n	Family	Correlation	Linear Span	Imbalance
Gold	$2m+1$	2^n+1	$1+2^{\frac{n+1}{2}}$	$2n$	$[1,2^{\frac{n+1}{2}}+1]$
Gold	$4m+2$	2^n-1	$1+2^{\frac{n+2}{2}}$	$2n$	$[1,2^{\frac{n+2}{2}}+1]$
Kasami (Small Set)	$2m$	$2^{\frac{n}{2}}$	$1+2^{\frac{n}{2}}$	$\frac{3n}{2}$	$[1,2^{\frac{n}{2}}+1]$
Kasami (Large Set)	$4m+2$	$2^{\frac{n}{2}}(2^n+1)$	$1+2^{\frac{n+2}{2}}$	$\leq \frac{5n}{2}$	$[1,2^{\frac{n+2}{2}}+1]$
Bent	$4m$	$2^{\frac{n}{2}}$	$1+2^{\frac{n}{2}}$	$\binom{n/2}{n/4}2^{\frac{n}{4}}$	1
No	$2m$	$2^{\frac{n}{2}}$	$1+2^{\frac{n}{2}}$	$m(2^m-1)$	$[1,2^{\frac{n}{2}}+1]$
TN	$2km$	$2^{\frac{n}{2}}$	$1+2^{\frac{n}{2}}$	$> 3mk(3k-1)^{m-2}$	$[1,2^{\frac{n}{2}}+1]$
$(0,j)$-QF	me	2^n+1	$1+2^{\frac{n+m}{2}}$	$\sim n2^{m-1}(e-2)^{m-2}$	$[1,2^{\frac{n+m}{2}}+1]$

Table 1: Comparison of Properties of Families of Sequences of Period 2^n-1

There is a spectrum of families of sequences exhibiting tradeoffs among the size of the family, the maximum cross-correlations, and the linear span. For families of sequences of period 2^n-1, with $2^{n/2}$ sequences in the family, various constructions have been described with successively greater linear spans. These include bent function sequences [6, 7, 11, 12], No sequences [9, 10], and TN sequences (also known as generalized No sequences) [4, 9]. The linear spans of these sequences are summarized in Table (1).

If we want larger families of sequences, we must weaken the bound on the maximum correlations. A number of such families have been studied, such as Gold sequences and large sets of Kasami sequences [1, 2], but the linear spans of these sequences are only linear in n.

In this paper we describe new families of binary sequences which we call *QF sequences* (QF stands for "quadratic form"). The size of these families is 2^n+1 when the period is 2^n-1. Certain of these sequences, $(0,j)$-QF sequences, have maximum correlations that are slightly greater than $2^{n/2}-1$, and linear spans which are larger than Gold, Kasami, and GMW sequences by a factor that is exponential in n. A parameter can be chosen that makes possible various combinations of maximum linear span and maximum correlation. Generally, the larger the linear span, the larger the maximum correlation that must be tollerated with these families. These results are also summarized in Table (1).

The construction of QF sequences is based on algebra over finite fields. They are special cases of the sequences known as d-form sequences [4]. The ith element of a QF sequence is given by starting with the ith power of a primitive element in the Galois field $GF(2^n)$, applying a quadratic form to $GF(2^m)$ (m a divisor of n), raising the result to some power, and mapping to $GF(2)$ by a trace function. We describe here the correlation and balance properties of $(0,j)$-QF sequences, based on previous results on TN sequences [4]. We also give lower bounds on the linear span of $(0,j)$-QF sequences, count the number of distinct families of such sequences, and discuss their implementation.

2 Definitions

Throughout this paper let e and m be positive integers, let $n = em$. For convenience we write $q = 2^m$. Let Tr_m^n be the trace function from $GF(2^n)$ to $GF(2^m)$,

$$Tr_m^n(x) = \sum_{i=0}^{e-1} x^{q^i}.$$

By a quadratic form on $GF(q)$ we mean a homogeneous polynomial of degree 2 in several variables.

By choosing coordinates, we can treat $GF(q^e)$ as e-dimensional affine space over $GF(q)$. Then any polynomial in e-variables with coefficients in $GF(q)$ defines a function from $GF(q^e)$ to $GF(q)$, and any such function is representable by a polynomial. Under this identification, the quadratic forms correspond precisely to sums of functions of the form

$$H(x) = Tr_m^n(\gamma x^t),$$

where the sum of the coefficients in the base q expansion of t is 2, $t < 1+2^{e/2}$, and $\gamma \in GF(q^e)$; and, if e is even, functions of the form

$$H(x) = Tr_m^{n/2}(\gamma x^{1+q^{e/2}}),$$

where $\gamma \in GF(q^{e/2})$.

QF sequences are generated in three steps. We start with a sequence of powers of a primitive element in $GF(q^e)$. To this sequence we apply a quadratic form H mapping to $GF(q)$. We then raise the result to some power. Finally, we apply the trace function Tr_2^m. The precise definition of QF sequences is as follows.

Definition 2.1 *Let e and m be positive integers, and let $q = 2^m$. Let r be a positive integer such that* $\gcd(r, q-1) = 1$. *Let α be a primitive element in $GF(q^e)$. Let $H(x)$ be a quadratic form on $GF(q^e)$ over $GF(q)$. Then the sequence* $\mathbf{S}$ *whose ith term is*

$$s_i = Tr_1^m((H(\alpha^i))^r) \tag{1}$$

is a QF sequence.

QF sequences are a special case of d-form sequences, in which the quadratic form H is replaced by a homogeneous polynomial of degree d [4].

The particular QF sequences we are interested in here are those of the form

$$s_i = Tr_1^m((Tr_m^n(\gamma\alpha^{2i}) + Tr_m^n(\delta\alpha^{(1+q^j)i}))^r).$$

We call such a sequence a $(0,j)$-*QF sequence*. We require that $\gcd(e,j)=1$. We sometimes also require that $\gcd(q^j+1, q^e-1)=1$. If $\gcd(e,j)=1$, this is equivalent to the condition that e is odd. We show that, for fixed n, m, j, and r, the family of all such $(0,j)$-QF sequences has near optimal cross-correlations and very high linear span. The number of cyclically distinct sequences in such a family is 2^n-1.

3 Cross-Correlations

In this section we determine the cross-correlations of the $(0,j)$-QF sequences defined in Section 2. Recall that the cross-correlation with shift τ of two sequences $\mathbf{S}=(s_1,s_2,\cdots)$ and $\mathbf{T}=(t_1,t_2,\cdots)$ of period N, is defined by

$$\Theta_{\mathbf{S},\mathbf{T}}(\tau) = \sum_{i=1}^{N}(-1)^{s_i+t_{i+\tau}}. \tag{2}$$

In our case $N=q^e-1$. The cross-correlations of two d-form sequences have been described in terms of the zeros of d-forms [4].

Theorem 3.1 (Klapper [4]) *Let the integers m, e, and r, and primitive element $\alpha \in GF(q^e)$ be fixed, and let H_1 and H_2 be d-forms on $GF(q^e)$ over $GF(q)$ defining d-form sequences $\mathbf{S}^1$ and $\mathbf{S}^2$. For any shift τ, let $z_\tau = |\{x \neq 0 \in GF(q^e) : H_1(x)+H_2(\alpha^\tau x)=0\}|$. Then*

$$\Theta_{\mathbf{S}^1,\mathbf{S}^2}(\tau) = \frac{qz_\tau - (q^e-1)}{q-1}.$$

In case $d=2$ and H is a quadratic form, the number of zeros of H is well understood. Recall that the *rank* of a quadratic form H is the smallest integer t such that there is a set of coordinates in which H can be represented using only t variables. The number of solutions x to the equation $H(x)=0$ (or, more generally, $H(x)=a$) is determined by the rank and, in the case of even rank, whether H is of one of two types. A nice treatment of this analysis can be found in Lidl and Niederreiter's book [8].

Let r, α, and j be fixed satisfying $\gcd(r,q-1)=1$. Consider a pair of $(0,j)$-QF sequences

$$s_i^k = Tr_1^m((Tr_m^n(\gamma_k\alpha^{2i}) + Tr_m^n(\delta_k\alpha^{i(1+q^j)}))^r),$$

where $\gamma_k, \delta_k \in GF(q^n)$, and $k = 1, 2$. To compute the cross-correlations of $\mathbf{S}^1$ and $\mathbf{S}^2$, we need only count the number z_τ of zeros of the quadratic form

$$Tr_m^n((\gamma_1 + \gamma_2\alpha^{2\tau})x^2) + Tr_m^n((\delta_1 + \delta_2\alpha^{(1+q^j)\tau})x^{1+q^j}) \overset{\text{def}}{=} Tr_m^n(\gamma x^2) + Tr_m^n(\delta x^{1+q^j}),$$

which we denote by $H(x)$.

If we could determine the rank of $H(x)$, we would be essentially done. Unfortunately, we have been unable to do so directly. However, the rank of the second summand has been analyzed [5], and we can use this analysis. We can determine z_τ from the numbers of solutions to certain systems of equations as follows. We have

$$Tr_m^n(\gamma x^2) + Tr_m^n(\delta x^{1+q^j}) = 0$$

if and only if for some $u \in GF(q)$

$$Tr_m^n(\gamma x^2) = u \text{ and } Tr_m^n(\delta x^{1+q^j}) = u. \tag{3}$$

Recall that every γ in $GF(q^n)$ can be written as a square, $\gamma = \beta^2$, with β in $GF(q^n)$, and that for any y, we have $Tr_m^n(y^2) = Tr_m^n(y)^2$. Letting $v^2 = u$, it follows that equation (3) holds if and only if

$$(Tr_m^n(\beta x))^2 = v^2 \text{ and } Tr_m^n(\delta x^{1+q^j}) = v^2,$$

which holds if and only if

$$Tr_m^n(\beta x) = v \text{ and } Tr_m^n(\delta x^{1+q^j}) = v^2. \tag{4}$$

Thus we can compute z_τ by solving a system consisting of a linear equation and a homogeneous quadratic equation. Such systems have been completely analyzed in terms of the rank and type of the quadratic form $Tr_m^n(\delta x^{1+q^j})$, and the rank and type of this quadratic form have been analyzed as well [5]. We quote the relevant results. In general, quadratic forms can be classified into three types. For simplicity, we write $\overline{x} = (x_1, \cdots, x_e)$. Let $B_t(\bar{x}) = x_1x_2 + x_3x_4 + \cdots + x_{t-1}x_t$. The function $\eta(v)$ is defined by

$$\eta(v) = \begin{cases} -1 & \text{if } v \neq 0 \\ q-1 & \text{if } v = 0. \end{cases}$$

Proposition 3.2 *Every quadratic form R of rank t in e variables over $GF(q)$, q even, is equivalent under a change of coordinates to one of the following three standard types:*

Type I: $B_t(\bar{x})$;

Type II: $B_{t-1}(\bar{x}) + x_t^2$;

Type III: $B_{t-2}(\bar{x}) + bx_{t-1}^2 + x_{t-1}x_t + bx_t^2$.

The number of solutions to the equation $R(\bar{x}) = v$ is:

Type I: $q^{e-1} + \eta(v)q^{e-t/2-1}$;

Type II: q^{e-1};

Type III: $q^{e-1} - \eta(v)q^{e-t/2-1}$.

For the quadratic forms in question, we have the following.

Theorem 3.3 (Klapper [5]) *Let $R(x) = Tr_m^n(\delta x^{1+q^j})$.*

1. *If $e/\gcd(e,j)$ is even and δ is not a $(1+q^j)$th power in $GF(q^e)$, then the rank of R is e, hence even. Moreover, if $e/(2\gcd(e,j))$ is odd, then R is a Type III quadratic form, while if $e/(2\gcd(e,j))$ is even, then R is a Type I quadratic form.*

2. *If $e/\gcd(e,j)$ is even and δ is a $(1+q^j)$th power in $GF(q^e)$, then the rank of R is $e - 2\gcd(e,j)$, hence even. Moreover, if $e/(2\gcd(e,j))$ is odd, then R is a Type I quadratic form, while if $e/(2\gcd(e,j))$ is even, then R is a Type III quadratic form.*

3. *If $e/\gcd(e,j)$ is odd, then the rank of R is $e - \gcd(e,j) + 1$, hence odd. Moreover, R is a Type II quadratic form.*

Thus if $\gcd(e,j) = 1$, then there are three possibilities:

1. Suppose $e \equiv 0 \pmod 4$. R has rank $e-2$ and Type III if δ is a $(1+q^j)$th power in $GF(q^e)$. Otherwise R has rank e and Type I.

2. Suppose $e \equiv 2 \pmod 4$. R has rank $e-2$ and Type I if δ is a $(1+q^j)$th power in $GF(q^e)$. Otherwise R has rank e and Type III.

3. Suppose e is odd. R has rank $e-1$ and Type II.

Furthermore, the solutions to systems of equations such as (4) have been determined. We first choose coordinates $x_1, \cdots, x_e$ so that $R(x)$ is one of the three standard types. Then we can write

$$Tr_m^n(\beta x) = \sum_{i=1}^{e} a_i x_i$$

for some $a_i \in GF(q)$. We let t be the rank of R and let $N(v)$ be the number of solutions to the system of equations (4). We also let $\epsilon = 1$ if R has even rank and Type I, and $\epsilon = -1$ if R has even rank and Type III. By applying Propositions 3.3 and 3.4 of [5], we have the following.

Proposition 3.4 *Suppose that for some $i > t$, $a_i \neq 0$.*

1. *If R has Type I or Type III, then*
$$N(v) = q^{e-2} + \epsilon\eta(v)q^{e-t/2-2}.$$

2. *If R has Type II then*
$$N(v) = q^{e-2}.$$

Proposition 3.5 *Suppose that $a_{t+1} = a_{t+2} = \cdots = a_e = 0$, or $t = e$.*

1. *Let R have Type I or Type III.*

   a. *If $R(\bar{a}) \neq 0$, then*
$$N(v) = \begin{cases} q^{e-2} + \epsilon(-1)^{\phi}q^{e-t/2-1} & \text{if } v \neq 0 \\ q^{e-2} & \text{if } v = 0, \end{cases}$$
   where $\phi = Tr_2^q(R(\bar{a}))$.

   b. *If $R(\bar{a}) = 0$, then*
$$N(v) = \begin{cases} q^{e-2} & \text{if } v \neq 0 \\ q^{e-2} + \epsilon(q-1)q^{e-t/2-1} & \text{if } v = 0. \end{cases}$$

2. *Let R have Type II.*

   a. *If $a_t = 0$, then $N(v) = q^{e-2}$.*

   b. *Otherwise $N(v) = q^{e-2} + (-1)^{\psi}\eta(v(1+a_t))q^{e-t/2-3/2}$, where $\psi = Tr_2^q(B_{t-1}(\bar{a})/a_t^2)$.*

These results can be combined with Theorems 3.1 and 3.3 to give the cross-correlations of quadratic form sequences. If $\mathbf{S}$ is a cyclic shift of $\mathbf{T}$, and τ is a shift so that $\mathbf{S}$ coincides with $\mathbf{T}$ shifted by τ, then the cross-correlation of $\mathbf{S}$ and $\mathbf{T}$ with shift τ is $2^n - 1$. We call this a *trivial* cross-correlation.

Theorem 3.6 *Let j and r be integers satisfying $\gcd(e, j) = \gcd(r, q-1) = 1$. Let $\mathbf{S}$ and $\mathbf{T}$ be two $(0, j)$-QF sequences with exponent r.*

1. *If e is even, then the nontrivial cross-correlations of $\mathbf{S}$ and $\mathbf{T}$ are $\{-1, \pm q^{e/2}-1, \pm q^{e/2+1}-1\}$.*

2. *If e is odd then the nontrivial cross-correlations of $\mathbf{S}$ and $\mathbf{T}$ are $\{-1, \pm q^{(e+1)/2} - 1\}$.*

Proof: The proof is completed by summing the appropriate entry in Proposition 3.4 or Proposition 3.5. In case 1 we have $t = e$ or $t = e - 2$. In case 2 we have $t = e$. We treat case 2 in detail and leave case 1 to the reader.

Since $t = e$, only Proposition 3.5 applies. If $a_e = 0$, then $N(v) = q^{e-2}$. Thus $z_\tau = \sum_v q^{e-2} - 1 = q^{e-1} - 1$, and

$$\Theta_{\mathbf{S},\mathbf{T}}(\tau) = \frac{qz_\tau - (q^e - 1)}{q-1} = -1.$$

If $a_e \neq 0$, then

$$N(v) = q^{e-2} + (-1)^\psi \eta(v(1+a_e))q^{(e-3)/2},$$

where $\psi = Tr_2^q(B_{e-1}(\bar{a})/a_e^2)$ is independent of v. If $a_e = 1$, then this reduces to

$$N(v) = q^{e-2} + (-1)^\psi (q-1)q^{(e-3)/2}.$$

Thus $z_\tau = q^{e-1} + (-1)^\psi (q-1)q^{(e-1)/2} - 1$, and

$$\begin{aligned}\Theta_{\mathbf{S},\mathbf{T}}(\tau) &= \frac{qz_\tau - (q^e - 1)}{q-1} \\ &= (-1)^\psi q^{(e+1)/2} - 1.\end{aligned}$$

If $a_e \notin \{0, 1\}$, then

$$N(v) = q^{e-2} + (-1)^\psi \eta(v)q^{(e-3)/2}.$$

Thus

$$\begin{aligned}z_\tau &= \sum_{v\neq 0}(q^{e-2} - (-1)^\psi q^{(e-3)/2}) + q^{e-2} + (-1)^\psi (q-1)q^{(e-3)/2} - 1, \\ &= q^{e-1} - 1,\end{aligned}$$

and $\Theta_{\mathbf{S},\mathbf{T}}(\tau) = -1$. □

The imbalance of a binary sequence is the number of times the sequence equals zero minus the number of times it equals one. The imbalances of $(0, j)$-QF sequences are in the same set of values as the cross-correlations.

Corollary 3.7 *Let j and r be integers satisfying* $\gcd(e, j) = \gcd(r, q - 1) = 1$. *Let* $\mathbf{S}$ *be a $(0, j)$-QF with exponent r.*

1. *If e is even, then the imbalance of* $\mathbf{S}$ *is in the set* $\{-1, \pm q^{e/2} - 1, \pm q^{e/2+1} - 1\}$.
2. *If e is odd then the the imbalance of* $\mathbf{S}$ *is in the set* $\{-1, \pm q^{(e+1)/2} - 1\}$.

We can also use Theorem 3.1 to show that the period of a quadratic form sequence is $2^n - 1$.

Theorem 3.8 *If* $\mathbf{S}$ *is a* $(0, j)$*-QF sequence based on the primitive element* $\alpha \in GF(2^n)$*, then the period of* $\mathbf{S}$ *is* $2^n - 1$*.*

Proof: Suppose $\mathbf{S}$ has period τ. Then the correlation of $\mathbf{S}$ with its own τ-shift (that is, the autocorrelation of $\mathbf{S}$ with shift τ) equals $2^n - 1$. By Theorem 3.1, this can only happen if $Tr^n_m(\gamma(1 + \alpha^{2\tau})x^2 + \delta(1 + \alpha^{\tau(1+q^j)})x^{1+q^j})$ is identically zero, where γ, δ, m, and j are the parameters defining $\mathbf{S}$. Therefore, $1 + \alpha^{2\tau} = 0$ or $1 + \alpha^{\tau(1+q^j)} = 0$ (one of γ and δ may be zero). But $2^n - 1$ is relatively prime to both 2 and $1 + q^j$, so $2^n - 1$ must be a divisor of τ. □

4 Linear Span

In this section we show that the linear span, $\lambda_{\mathbf{S}}$, of a $(0, j)$-QF sequence $\mathbf{S}$ is at least that of a GMW sequence with the same period, and that when γ and δ are nonzero, the linear span asymptotically exceeds the linear span of a GMW sequence by a factor of $2^{\mathrm{wt}(r)}$, where $\mathrm{wt}(r)$ is the number of ones in the binary expansion of r. If $\gamma = 0$ or $\delta = 0$, then $\mathbf{S}$ is a GMW sequence and, as is well known, the linear span is $me^{\mathrm{wt}(r)}$.

The binary expansion of the exponent r in the definition of a $(0, j)$-QF sequence is a series of runs of ones separated by zeros. We let U be the number of runs of ones and let R_i be the length of the ith run.

Theorem 4.1 *Let* $\mathbf{S}$ *be a* $(0, j)$*-QF sequence with*

$$s_i = Tr^m_1((Tr^n_m(\gamma\alpha^i)^2 + Tr^n_m(\delta\alpha^{i(1+q^j)}))^r).$$

Suppose γ *and* δ *are nonzero. Let* $\pi = (e^2 - 4e + 1)^{1/2}$*. If* $e > 3$*, then the linear span of* $\mathbf{S}$ *is at least*

$$m \prod_{i=1}^{U} \left(\frac{(e + 1 + \pi)(e - 1 + \pi)^{R_i} - (e - 1 - \pi)(e + 1 - \pi)^{R_i}}{2\pi} \right)$$

$$\geq \begin{cases} 2n \cdot 4^{\mathrm{wt}(r)-1} & \text{if } e = 4 \\ n \cdot \prod_{i=1}^{U}(2^{R_i}(e - 2)^{R_i - 1} - 2 \cdot 4^{R_i - 1}) & \text{otherwise.} \end{cases}$$

This last product is maximized when $r = 2^{m-1} - 1$ *at* $n(2^{m-1}(e - 2)^{m-2} - 2 \cdot 4^{m-2})$*. If* $e = 3$*, then the linear span of* $\mathbf{S}$ *is at least*

$$m6^U \cdot 4^{\mathrm{wt}(r)}.$$

This is maximized at $m \cdot 6^{m/2} \cdot 4^{m/2}$ *when* m *is even by taking* $r = 1 + 2^2 + 2^4 + \cdots 2^{m-2}$*. It is maximized at* $m \cdot 6^{(m-1)/2} \cdot 4^{(m+1)/2}$ *when* m *is odd by taking* $r = 1 + 2 + 2^3 + 2^5 + \cdots 2^{m-2}$*.*

Proof: Key [3] showed that if we express the ith term of a sequence $\mathbf{S}$ as a polynomial in α^i, then $\lambda_{\mathbf{S}}$ is the number of monomials in that polynomial. Equivalently, the linear span is the number of monomials in the polynomial

$$s(x) \quad = \quad Tr_1^m((Tr_m^n(\gamma x)^2 + Tr_m^n(\delta x^{1+q^j}))^r).$$

We show that, when this expression is expanded to a sum of monomials, all but a small number of monomials are distinct. We explicitly describe those monomials that coincide with others, and count the remaining monomials.

Let the binary expansion of r be $r = \sum_{\ell=0}^{m-1} a_\ell 2^\ell$, $a_\ell \in \{0,1\}$. Let $L = \{\ell : a_\ell \neq 0\}$. Let Δ be the set of wt(r)-tuples of elements of $\{0,\cdots,e-1\}$, indexed by L, and let Γ be the the set of wt(r)-tuples of elements of $\{2, 1+q^j\}$, indexed by L. For $I \in \Delta$ and $K \in \Gamma$, we write i_ℓ and k_ℓ for the ℓth components of I and K, respectively.

Lemma 4.2 *The polynomial $s(x)$ can be expanded to a sum of monomials of the form $x^{\mu(t,I,K)}$, where*

$$\mu(t,I,K) = 2^t \sum_{\ell \in L} 2^\ell k_\ell q^{i_\ell},$$

$0 \leq t < m$, I ranges over Δ, and K ranges over Γ.

Proof: We have

$$\begin{aligned}
s(x) &= Tr_1^m((Tr_m^n(\gamma x)^2 + Tr_m^n(\delta x^{1+q^j}))^r) \\
&= Tr_1^m((\sum_{i=0}^{e-1} \gamma^{2q^i} x^{2q^i} + \sum_{i=0}^{e-1} \delta^{q^i} x^{q^i+q^{j+i}})^r) \\
&= Tr_1^m(\prod_{\ell \in L} \sum_{i=0}^{e-1} (\gamma^{2^{\ell+1}q^i} x^{2^{\ell+1}q^i} + \delta^{2^\ell q^i} x^{2^\ell(q^i+q^{j+i})})) \\
&= Tr_1^m(\sum_{I \in \Delta} \prod_{\ell \in L} (\gamma^{2^{\ell+1}q^{i_\ell}} x^{2^{\ell+1}q^{i_\ell}} + \delta^{2^\ell q^{i_\ell}} x^{2^\ell(q^{i_\ell}+q^{j+i_\ell})})) \\
&= Tr_1^m(\sum_{I \in \Delta} \sum_{K \in \Gamma} c_{I,K} \prod_{\ell \in L} x^{2^\ell k_\ell q^{i_\ell}}) \\
&= Tr_1^m(\sum_{I \in \Delta} \sum_{K \in \Gamma} c_{I,K} x^{\sum_{\ell \in L} 2^\ell k_\ell q^{i_\ell}}) \\
&= \sum_{t=0}^{m-1} \sum_{I \in \Delta} \sum_{K \in \Gamma} d_{t,I,K} x^{2^t \sum_{\ell \in L} 2^\ell k_\ell q^{i_\ell}},
\end{aligned}$$

where $c_{I,K}$ and $d_{t,I,K}$ are nonzero elements of $GF(q^e)$. Thus the exponent of the general term is of the form

$$\mu(t,I,K) = 2^t \sum_{\ell \in L} 2^\ell k_\ell q^{i_\ell}.$$

□

We use the following terminology. Each $\mu(t,I,K)$ is an *exponent*. It is a sum of *terms* $2^\ell k_\ell q^{i_\ell}$, with $k_\ell \in \{2, 1+q^j\}$. Each term is a sum of one or two *summands*, $2^{\ell+1}q^{i_\ell}$; or $2^\ell q^{i_\ell}$ and $2^\ell q^{i_\ell+j}$.

Let $\mu(t,I,K)$ and $\mu(t',I',K')$ be two exponents. We say (t,I,K) is *equivalent* to (t',I',K'), written $(t,I,K) \sim (t',I',K')$, if

$$\mu(t,I,K) \equiv \mu(t',I',K') \pmod{q)^e - 1}. \tag{5}$$

Two index triples are equivalent if and only if they correspond to the same powers of x, but with possibly different coefficients. Their resulting monomials may (or may not) cancel. Thus we want to count the number of (t,I,K) that are equivalent to other index triples, and subtract this number from the total number of index triples. This gives us a lower bound on the number of monomials in $s(x)$ with nonzero coefficients, and hence a lower bound on $\lambda_{\mathbf{S}}$.

If we reduce equation (5) modulo $q-1$, then it becomes to $2^{t+1}r \equiv 2^{t'+1}r \pmod{q) - 1}$. Since $\gcd(r,q-1) = 1$ and $q-1$ is odd, it follows that $2^{t-t'} \equiv 1 \pmod{q) - 1} = 2^m - 1$. Thus m divides $t - t'$. But t and t' are between 0 and $m-1$, so we must have $t = t'$. Thus it suffices to consider the case when $t = t' = 0$. From now on we drop t from the notation and simply write $\mu(I,K)$. The following lemma is useful.

Lemma 4.3 *Let J and J' be index sets, and, for each $j \in J$ (respectively, $j \in J'$) let i_j (respectively, i'_j) be a nonnegative integer. Suppose that*

$$\sum_{j\in J} 2^{i_j} = \sum_{j\in J'} 2^{i'_j} \tag{6}$$

and that the i'_j are distinct. Then there is a partition of J into disjoint sets, $J = \cup_{u\in J'} J_u$ such that

$$\sum_{j\in J_u} 2^{i_j} = 2^{i'_u}.$$

Proof: The proof is by induction on $|J|$. If $|J| = 1$, then $|J'| = 1$, and the result is trivial.

Let $|J| > 1$ and let i_{j_1} be the minimal element of $\{i_j : j \in J\}$. If there is a $j_0 \in J'$ such that $i'_{j_0} = i_{j_1}$, then we let $J_{j_0} = \{i_{j_1}\}$. By induction, we can partition the remainder of J as required.

If there is no such $j_0 \in J'$, then $i'_j > i_{j_1}$ for every $j \in J'$. The right hand side of equation (6) is divisible by $2^{i_{j_1}+1}$, so a positive even number of the i_j equal i_{j_1}. Let $j_2 \in J$ satisfy $i_{j_2} = i_{j_1}$. Let d be a new index not in J, and let $i_d = i_{j_1} + 1$. Then $J - \{j_1, j_2\} \cup \{d\}$ and J' satisfy the hypotheses of the lemma, so by induction the conclusions hold for this set of indices. By replacing d by j_1 and j_2 in the partition containing d, we get a partition of J satisfying the lemma. □

It follows that for any $\mu(I,K)$, either all summands are distinct, or there is a subset of the summands that forms a partition as in the lemma. Let J be such a partition, with $\sum_{j\in J} 2^{i_j} = 2^d$, and $|J| \neq 1$. Let i_{j_1} be the least element in $\{i_j\}$. There must be a second $j_2 \in J$ with $i_{j_2} = i_{j_1}$. The term in $\mu(I,K)$ contributing i_{j_1} is of the form $2^\ell x q^i$, with $x \in \{1,2,q^j\}$, and the term in $\mu(I,K)$ contributing i_{j_2} is of the form $2^{\ell'} y q^{i'}$, with $y \in \{1,2,q^j\}$.

We must have $\ell \neq \ell'$. Reducing modulo $q-1$, it is apparent that ℓ and ℓ' must differ by 1, and x or y (whichever corresponds to the smaller of ℓ and ℓ') equals 2. Thus we may assume $x = 2$, $y \in \{1, q^j\}$, and $\ell' = \ell+1$. If $y = 1$, then $i' = i$, while if $y = q^j$, then $i' = i-j \pmod{e}$. Furthermore, these are the only terms contributing summands whose exponents are congruent to $\ell+1$ modulo m. Thus the two terms contributing these summands are

$$2^\ell \cdot 2 \cdot q^i + 2^{\ell+1}(1+q^j)q^i = 2^\ell \cdot 2 \cdot q^{i+j} + 2^{\ell+1} \cdot 2 \cdot q^i$$

or

$$2^\ell \cdot 2 \cdot q^i + 2^{\ell+1}(1+q^j)q^{i-j} = 2^\ell \cdot 2 \cdot q^{i-j} + 2^{\ell+1} \cdot 2 \cdot q^i.$$

In either case, replacing the terms on the left hand side by the terms on the right hand side results in an exponent $\mu(I',K') = \mu(I,K)$ with $(I',K') \neq (I,K)$, and the new exponent has fewer summands.

Thus any exponent with a nontrivial partition of its summands, as in Lemma 4.3, is equivalent to another exponent, and must contain consecutive terms with $k_\ell = 2$, $k_{\ell+1} = 1+q^j$, and $i_{\ell+1} \in \{i_\ell, i_\ell - j\}$. Any exponent that is equivalent to another exponent with fewer summands must contain a notrivial partition of its summands. Conversely, any exponent with consecutive terms with $k_\ell = k_{\ell+1} = 2$, and $i_{\ell+1} \in \{i_\ell - j, i_\ell + j\}$ is equivalent to another exponent with more summands. Any exponent that is equivalent to another exponent with more summands must have terms of this form.

Thus if we exclude all exponents with consecutive terms of these four types, the only remaining equivalences are between exponents with no nontrivial partitions of their summands. Assume $\mu(I,K)$ and $\mu(I',K')$ is such a pair of distinct equivalent exponents. The set of summands of $\mu(I',K')$ must be a permutation of the set of summands of $\mu(I,K)$.

Since these exponents are distinct, there is at least one ℓ for which $k_\ell = 2$ and either $k_\ell \neq k'_\ell$ or $i_\ell \neq i'_\ell$. Fix this ℓ and suppose $k'_\ell = 1 + q^j$. Then $\mu(I,K)$ has at most one summand of the form $2^\ell q^i$ (arising from the term with index $\ell - 1$), but $\mu(I',K')$ has two such summands. Thus it must be that $k'_\ell = 2$, and so $i_\ell \neq i'_\ell$.

The term with index ℓ in each exponent has the form $2^{\ell+1}q^i$, and these terms are distinct, so each exponent must have other summands with this form. The only other terms such summands can come from are those with index $\ell+1$, and so we must have $k_{\ell+1} = k'_{\ell+1} = 1+q^j$.

Thus in each exponent there is a set of three summands of the form $2^{\ell+1}q^i$, and these sets are permutations of each other. The only possibilities are

$$\mu(I,K) = 2^{\ell+1}q^i + 2^{\ell+1}(1+q^j)q^{i+j} = 2^{\ell+1}q^{i+2j} + 2^{\ell+2}(1+q^j)q^i = \mu(I',K')$$

and the reverse.

If we exclude all exponents with consecutive terms of these two types as well as the preceding four types, we will have excluded exactly those exponents that are equivalent to other exponents.

Proposition 4.4 *The only exponents that are equivalent to other exponents are those containing consecutive terms of one of the following six types:*

1. $k_\ell = 2$, $k_{\ell+1} = 1+q^j$, $i_{\ell+1} = i_\ell$;
2. $k_\ell = 2$, $k_{\ell+1} = 1+q^j$, $i_{\ell+1} = i_\ell - j$;
3. $k_\ell = 2$, $k_{\ell+1} = 2$, $i_{\ell+1} = i_\ell - j$;
4. $k_\ell = 2$, $k_{\ell+1} = 2$, $i_{\ell+1} = i_\ell + j$;
5. $k_\ell = 2$, $k_{\ell+1} = 1+q^j$, $i_{\ell+1} = i_\ell + j$;
6. $k_\ell = 2$, $k_{\ell+1} = 1+q^j$, $i_{\ell+1} = i_\ell - 2j$.

Thus the linear span of **S** *is at least the number of exponents with no such consecutive terms of these types. We call these* good *exponents.*

Note that if $e = 3$, then $i_\ell - 2j \equiv i_\ell + j \pmod{e}$, so the fifth and sixth types coincide and there is no distinct pair of equivalent exponents. Otherwise all six types are distinct.

Any exponent for which all $k_\ell = 1+q^j$ is good. Thus there are at least $e^{\mathrm{wt}(r)}$ good exponents for each t, $0 \le t < m$, and the linear span is at least $me^{\mathrm{wt}(r)}$.

Consider a single run of ones of length R in the binary expansion of r. We can count the number of (i_ℓ, k_ℓ) for ℓ in this run that give good exponents by building them up a term at a time, from small ℓ to large ℓ. Given an exponent that is good up to its ℓth term, we count the ways it can be extended to a good exponent by adding a term for the next bit of r. If the next bit of r is the first in a run of ones, or if $k_\ell = 1+q^j$, then there are no restrictions, so there are $2e$ ways of extending. Otherwise, there are $e-4$ ways of extending the exponent with $k_{\ell+1} = 1+q^j$ (one way if $e = 3$), and $e-2$ ways of extending it with $k_{\ell+1} = 2$.

It follows that the good choices of $\{k_\ell\}$ and $\{i_\ell\}$ for one run are independent of those for other runs. The total number of good exponents is therefore the product of the number of good choices for the different runs. For each run, we can count the good choices by a recurrence. Only the length of the run matters, so we can assume the run is a sequence of consecutive indices $0, 1, 2, \cdots, R-1$. We let $A(\ell)$ be the number of good runs up to index ℓ, ending with $k_\ell = 2$, and let $B(\ell)$ be the number of good runs up to index ℓ, ending with $k_\ell = 1 + q^j$. Then, for $e > 3$, we have the system of recurrences

$$\begin{aligned} A(\ell) &= (e-2)A(\ell-1) + eB(\ell-1) \\ B(\ell) &= (e-4)A(\ell-1) + eB(\ell-1), \end{aligned} \tag{7}$$

for $1 \le \ell < R$. For $e = 3$, we have the system

$$\begin{aligned} A(\ell) &= A(\ell-1) + 3B(\ell-1) \\ B(\ell) &= A(\ell-1) + 3B(\ell-1), \end{aligned}$$

for $1 \le \ell < R$. In both cases the initial conditions are $A(0) = B(0) = e$.

First consider the case when $e = 3$. We have $A(\ell) = B(\ell)$ for all ℓ, so $A(\ell) = 4A(\ell-1)$. The solution to this recurrence is $A(\ell) = 3 \cdot 4^\ell$, so $A(\ell) + B(\ell) = 6 \cdot 4^\ell$. Thus the total number of good exponents for a run of length R when $e = 3$ is $6 \cdot 4^R$. The total number of good exponents for a given r is the product

$$\prod_{i=1}^{U} 6 \cdot 4^{R_i} = 6^U \cdot 4^{\mathrm{wt}(r)},$$

where i ranges over the set of runs of ones in the binary expansion of r, R_i is the length of the ith run, and U is the number of runs. For each good exponent $\mu(I, K)$ and each t, we have a monomial

$$x^{2^t \sum_{\ell \in L} 2^\ell k_\ell q^{i_\ell}}$$

that must appear in $s(x)$ with a nonzero coefficient. Thus the linear span is at least

$$m6^U \cdot 4^{\mathrm{wt}(r)}.$$

This is maximized at $m \cdot 6^{m/2} \cdot 4^{m/2}$ when m is even by taking $r = 1 + 2^2 + 2^4 + \cdots 2^{m-2}$. It is maximized at $m \cdot 6^{(m-1)/2} \cdot 4^{(m+1)/2}$ when m is odd by taking $r = 1 + 2 + 2^3 + 2^5 + \cdots 2^{m-2}$.

The general case is somewhat more complicated. Equation (7) can be interpretted as a matrix equation,

$$\begin{pmatrix} A(\ell) \\ B(\ell) \end{pmatrix} = \begin{pmatrix} e-2 & e \\ e-4 & e \end{pmatrix} \cdot \begin{pmatrix} A(\ell-1) \\ B(\ell-1) \end{pmatrix}. \tag{8}$$

Let M be the matrix

$$M = \begin{pmatrix} e-2 & e \\ e-4 & e \end{pmatrix}.$$

Equation (8) is solved by diagonalizing M. Let $\pi = (e^2 - 4e + 1)^{1/2}$. Then $M = D^{-1}ND$, where

$$D = \begin{pmatrix} -1+\pi & e \\ -1-\pi & e \end{pmatrix},$$

and

$$N = \begin{pmatrix} e+\pi-1 & 0 \\ 0 & e-\pi+1 \end{pmatrix}.$$

The total number of good exponents for a run of length R is therefore given by

$$\begin{aligned} \begin{pmatrix} 1 & 1 \end{pmatrix} M^{R-1} \begin{pmatrix} e \\ e \end{pmatrix} &= \begin{pmatrix} 1 & 1 \end{pmatrix} D^{-1} N^{R-1} D \begin{pmatrix} e \\ e \end{pmatrix} \\ &= \left(\frac{e+1+\pi}{2\pi}(e-1+\pi)^R - \frac{e-1-\pi}{2\pi}(e+1-\pi)^R \right) \end{aligned}$$

The total number of good exponents for a given r is therefore the product

$$\prod_i \left(\frac{e+1+\pi}{2\pi}(e-1+\pi)^{R_i} - \frac{e-1-\pi}{2\pi}(e+1-\pi)^{R_i} \right).$$

This is maximized when $\mathrm{wt}(r) = m-1$, and r has a single run of ones. In this case the linear span is at least

$$\left(\frac{e+1+\pi}{2\pi}(e-1+\pi)^{m-1} - \frac{e-1-\pi}{2\pi}(e+1-\pi)^{m-1} \right).$$

The estimates in the statement of the theorem follow from the observations that $e-3 < \pi < e-2$ when $e \geq 5$, and that $\pi = 1$ when $e = 4$. □

Note that the linear span may be larger than these estimates. Each bad exponent is in a family of two or more exponents that contribute to the same term in the final polynomial expansion. Each of these exponents contributes a coefficient. If these sum to zero, then the linear span is unaffected by this term. Otherwise, it is increased by one. The coefficient of such a term is a sum of monomials in γ and δ (when $k_\ell = 2$, γ occurs with exponent $2^{\ell+1}q^{i_\ell}$, otherwise δ occurs with exponent $2^\ell q^{i_\ell}$). These expressions appear quite difficult to analyze, though it is conceivable that one could upper bound the number that can vanish.

5 Implementation of $(0, j)$-QF Sequence Generators

Consider the $(0, j)$-QF sequence $\mathbf{S}$ with

$$\begin{aligned} s_i &= Tr_1^m((Tr_m^n(\gamma\alpha^{2i} + \delta\alpha^{i(1+2^{mj})}))^r) \\ &= Tr_1^m((Tr_m^n(\gamma\alpha^{2i}) + Tr_m^n(\delta\alpha^{i(1+2^{mj})}))^r) \end{aligned}$$

Since α is a primitive element in $GF(2^n)$, α^2 and $\alpha^{1+2^{mj}}$ are also primitive elements in $GF(2^n)$. It follows that the sequences

$$Tr_m^n(\alpha^{2i})$$

and

$$Tr_m^n(\alpha^{i(1+2^{mj})})$$

are m-sequence of elements in $GF(2^m)$, and can be generated by linear feedback shift registers (LFSRs) of length n/m over $GF(2^m)$. That is, the elements of the registers are elements of $GF(2^m)$, and the feedback functions are linear functions in n/m variables. Such a register requires only n bits of storage. The arithmetic required is at most n/m multiplications by constants in $GF(2^m)$ (the coefficients of the minimum polynomial of α^2 or $\alpha^{1+2^{mj}}$, some of which may be zero), and at most $n/m - 1$ additions in $GF(2^m)$. The arithmetic can be minimized if α is chosen to minimize the number of nonzero coefficients in the minimal polynomials of α^2 and $\alpha^{1+2^{mj}}$ over $GF(2^m)$.

One extra addition is required to combine the outputs of the two LFSRs. The result is then raised to the rth power, and the trace to $GF(2)$ computed. However, in representing elements of $GF(2^m)$ as m-bit vectors, we can choose a basis so that the trace of an element is always given by projection onto a fixed component, say the first. Thus we only need to compute a single bit of the rth power.

The different choices of γ and δ correspond to different initial loadings of the LFSRs. Thus an entire family of TN sequences can be implemented by a single hardware circuit. Changing to a new sequence is possible by simply resetting the initial loading of the LFSRs.

6 The Number of Distinct Families of $(0, J)$-QF Sequences

It is useful to know how many distinct families of $(0, j)$-QF sequences there are with a given period $2^n - 1$. In this section we keep n fixed and let the factorization $n = me$ vary. We show that each choice of parameters m (dividing n, with $e = n/m$), j (relatively prime to e and less than $e/2$), r (up to multiplication by a power of 2), and α (up to raising to an exponent which is a power of two) gives rise to a distinct family of sequences, in the sense that no sequence in one family is a cyclic shift of a sequence in another family. For any fixed even integer n we write

$$S(m,j,\alpha,r) = \{\mathbf{S}^{\gamma,\delta} : s_i = Tr_1^m((Tr_m^n(\gamma\alpha^{2i} + \delta\alpha^{i(1+2^{mj})}))^r) \text{ and } \gamma,\delta \in GF(2^n)\},$$

where m divides n, $e = n/m$, $\gcd(j,e) = 1$, $j < e/2$, α is a primitive element of $GF(2^n)$, and r is relatively prime to $2^m - 1$. The family does not change when r is multiplied by a power of two, so we can assume that $r < 2^{m-1}$ and is odd. Note that a cyclic shift of one of these sequences by τ places replaces γ by $\gamma\alpha^{2\tau}$, and replaces δ by $\delta\alpha^{\tau(1+2^{mj})}$, and hence gives another sequence in the family.

Proposition 6.1 *Let n be a positive integer, $N = 2^n - 1$, m_1, and m_2 be divisors of n such that m_1 and m_2 divide n, and $e_1 = n/m_1$ and $e_2 = n/m_2$. Let r_1, r_2, j_1, and j_2 be integers such that $1 \le r_k < 2^{m_k - 1}$, r_k is relatively prime to $2^{m_k} - 1$, $j_k < e_k/2$, and j_k is relatively prime to e_k for $k = 1, 2$. Let α_1 and α_2 be primitive elements in $GF(2^n)$. Then $S(m_1, j_1, \alpha_1, r_1)$ and $S(m_2, j_2, \alpha_2, r_2)$ are disjoint unless either*

1. *$m_1 = m_2$, $j_1 = j_2$, and for some integers u and v, $0 \le u < n$, and $0 \le v < m_1$, $\alpha_2 = \alpha_1^{2^u}$, and $r_1 = 2^v \cdot r_2$, or*

2. *$m_1 j_1 = m_2 j_2$, r_1 and r_2 are powers of 2, and for some integer u, $0 \le u < n$ and $\alpha_2 = \alpha_1^{2^u}$.*

In each of these cases

$$S(m_1, \alpha_1, r_1) = S(m_2, \alpha_2, r_2).$$

Proof: Suppose that we have a pair of sequences in $\mathbf{S}^1 = S(m_1, j_1, \alpha_1, r_1)$ and $\mathbf{S}^2 = S(m_2, j_2, \alpha_2, r_2)$ respectively, such that one is a cyclic shift of the other. Since shifting a sequence gives another one in the same family, we can assume $\mathbf{S}^1 = \mathbf{S}^2$.

If r_1 and r_2 are powers of two, they can be factored out of the trace functions. Thus in this case we can assume $r_1 = r_2 = 1$, and $m_1 = m_2 = 1$. For some k, we have $\alpha_2 = \alpha_1^k$, so for every $x \in GF(2^n)$,

$$Tr_1^n(\gamma_1 x^2 + \delta_1 x^{1+2^{j_1}}) = Tr_1^n(\gamma_2 x^{2k} + \delta_2 x^{k(1+2^{j_2})}).$$

Expanding the traces and comparing the degrees of the terms on each side of the equality, we see that k must be a power of two, and $j_1 = j_2$. Furthermore, if only one of r_1 and r_2 is power of two, then the two sequences have different linear span, which is impossible.

In what remains we assume that neither r_1 nor r_2 is a power of two. Since n is fixed, there are integers a, b, c, and d with $abcd = n$ and b relatively prime to c, such that $m_1 = ab$

and $m_2 = ac$. By symmetry we can assume $b < c$. Let L_v be the set of indices of nonzero coefficients in the binary expansion of r_v, $v = 1, 2$.

The sequences $\mathbf{S}^1$ and $\mathbf{S}^2$ are equal if and only if the associated polynomials $s^1(x)$ and $s^2(x)$ are equal. When we expand $s^1(x)$ and $s^2(x)$ as sums of monomials, the typical term in $s^1(x)$ has exponent $\mu_1(t, I, K) = 2^t \sum_{\ell \in L_1} 2^\ell k_\ell 2^{abi_\ell}$ with $0 \le t < ab$, $k_\ell \in \{2, 1 + 2^{abj_1}\}$, and $0 \le i_\ell < cd$. The typical term in $s^2(x)$ has exponent $\mu_2(t, I, K) = 2^t \sum_{\ell \in L_2} 2^\ell k_\ell 2^{aci_\ell}$, with $0 \le t < ac$, $k_\ell \in \{2, 1 + 2^{acj_1}\}$, and $0 \le i_\ell < bd$. For some choices of parameters there may be cancellation, but, as we have seen, if all $k_\ell = 1 + 2^{abj_1}$ in the first case, and all $k_\ell = 1 + 2^{acj_1}$ in the second case, then the terms with these exponents are not cancelled. Moreover, these exponents are the only ones whose base two expansions have the maximum weight in each $s^v(x)$. Thus their weights, $2 \cdot \mathrm{wt}(r_1)$ and $2 \cdot \mathrm{wt}(r_2)$ must be equal. Hence $\mathrm{wt}(r_1) = \mathrm{wt}(r_2)$, and the set of these exponents for $s^1(x)$ is a permutation of the set of these exponents for $s^2(x)$.

In particular, there are t and $I = \langle i_\ell \rangle$ such that

$$\sum_{\ell \in L_1} 2^\ell (1 + 2^{abj_1}) = 2^t \sum_{\ell \in L_2} 2^\ell (1 + 2^{acj_2}) 2^{aci_\ell}.$$

We can simplify this by replacing r_2 by $2^t r_2$. This has no effect on the sequences, but changes L_2. Thus we can assume that $t = 0$, so

$$\sum_{\ell \in L_1} 2^\ell (1 + 2^{abj_1}) = \sum_{\ell \in L_2} 2^\ell (1 + 2^{acj_2}) 2^{aci_\ell}.$$

This implies that for each $\ell \in L_1$ there is an $\ell' \in L_2$ such that one of the two following cases holds.

1. $\ell = \ell' + aci_{\ell'}$: Since $0 \le \ell < ab < ac$, and $0 \le \ell' < ac$, this is only possible if $i_{\ell'} = 0$ and $\ell = \ell'$. Also, we must have $\ell + acj_2 = \ell'' + abj_1$ for some $\ell'' \in L_1$. Thus
$$b - 1 \ge |(\ell - \ell'')/a| = |bj_1 - cj_2|. \tag{9}$$

2. $\ell \equiv \ell' + aci_{\ell'} + acj_2 \pmod{a}bcd$: This implies $\ell \equiv \ell' \pmod{a}c$. But $0 \le \ell < ac$ and $0 \le \ell' < ac$, so $\ell = \ell'$. We must have $i_{\ell'} = bd - j_2 > bd/2$. Then $\ell + aci_{\ell'} = \ell + acb - j_2 = \ell'' + abj_1$ for some $\ell'' \in L_1$. As in case (1), it follows that
$$b - 1 \ge |c(bd - j_2) - bj_1|. \tag{10}$$

Note that for each $\ell \in L_1$, whichever case holds we have $\ell \in L_2$, i.e., $L_1 \subseteq L_2$. These sets have the same cardinality, so $L_1 = L_2 \overset{def}{=} L$, which means that $r_1 = r_2 \overset{def}{=} r$. Moreover, equations (9) and (10) together imply that $c(bd - 2j_2) \le 2(b - 1)$. But $j_2 \le bd/2 - 1$, so

$2c \leq c(bd-2j_2) \leq 2(b-1)$, i.e., $c < b$, contrary to our assumption. Therefore either case (1) holds for every ℓ, or case (2) holds for every ℓ.

If case (1) holds, then $\sum_{\ell\in L} 2^\ell(1+2^{abj_1}) = \sum_{\ell\in L} 2^\ell(1+2^{acj_2})$ so that $\sum_{\ell\in L} 2^{\ell+abj_1} = \sum_{\ell\in L} 2^{\ell+acj_2}$. Each $\ell+abj_1$ and each $\ell+acj_2$ is less than $abcd$, so the sequence of $\ell + abj_1 \pmod{a}bcd$ increases monotonically with ℓ. The same holds for the sequence of $\ell+acj_2$. The only possibility is that, for each ℓ, $2^{\ell+abj_1} = 2^{\ell+acj_2}$, so $abj_1 = acj_2$. The relative primality of b and c implies that there is a $j < d$ with $j_1 = cj$, and $j_2 = bj$. A similiar argument in case (2) leads to the equation $abj_1 = abcd - acj_2$. This is impossible, since $j_1 < cd/2$ and $j_2 < bd/2$.

It follows that

$$\begin{aligned} s^1(x) &= Tr_1^{ab}((Tr_{ab}^n(\gamma_1 x^2 + \delta_1 x^{1+2^{abcj}}))^r) \\ &= Tr_1^{ac}((Tr_{ac}^n(\gamma_2 x^2 + \delta_2 x^{1+2^{abcj}}))^r) \\ &= s^2(x). \end{aligned} \tag{11}$$

Let ψ be an element of $GF(2^{abcd})$ such that

$$Tr_{abc}^n(\gamma_1\psi^2 + \delta_1\psi^{1+2^{abcj}}) \stackrel{def}{=} \beta_1 \neq 0,$$

and let

$$Tr_{abc}^n(\gamma_2\psi^2 + \delta_2\psi^{1+2^{abcj}}) \stackrel{def}{=} \beta_2.$$

Such a ψ must exist since $Tr_{abc}^n(\gamma_1 x^2 + \delta_1 x^{1+2^{abcj}})$ is a nonzero quadratic form. If we restrict x to be of the form $x = \psi y$, with $y \in GF(2^{abc})$ then equation (11) reduces to

$$Tr_1^{ab}((Tr_{ab}^{abc}(\beta_1 x^2)^r) = Tr_1^{ac}((Tr_{ac}^{abc}(\beta_2 x^2))^r). \tag{12}$$

Since β_1 is nonzero, β_2 must also be nonzero. However, the sequence associated with the left hand side of equation (12) has linear span $ab \cdot c^{\mathrm{wt}(r)}$, while the sequence associated with the right hand side of equation (12) has linear span $ac \cdot b^{\mathrm{wt}(r)}$. Since these are equal, the weight of r must be one. That is, r must be a power of two. □

It is now straightforward to determine the number of distinct families of $(0,j)$-QF sequences.

Theorem 6.2 *Let n be a nonnegative integer. The number $\mathbf{N}_{\mathrm{QF}}$ of distinct families of $(0,j)$-QF sequences of period 2^n-1 is given by*

$$\mathbf{N}_{\mathrm{QF}} = \frac{\phi(2^n-1)}{n} \cdot \left(\sum_{e|n, e>1} \frac{\phi(e)}{2} \left(\frac{e\phi(2^{n/e}-1)}{n} - 1 \right) + \frac{\phi(n)}{2} \right),$$

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

If n is not a power of 2, The number $\mathbf{N}'_{\rm QF}$ of distinct families of $(0,j)$-QF sequences of period 2^n-1 with correlations in $\{\pm q^{(e+1)/2}-1,-1\}$ is given by

$$\mathbf{N}'_{\rm QF}=\frac{\phi(2^n-1)}{n}\cdot\left(\sum_{\substack{e|n\\ e>1\ odd}}\frac{\phi(e)}{2}\left(\frac{e\phi(2^{n/e}-1)}{n}-1\right)+\frac{\phi(n')}{2}\right),$$

where n' is the maximal odd divisor of n.

Proof: If α_1 and α_2 are primitive elements in $GF(2^n)$, then they are equivalent for purposes of generating families of $(0,j)$-QF sequences if they are in the same Galois coset, i.e., if $\alpha_2=\alpha_1^{2^k}$ for some k. If $e|n$ is chosen with e odd, and $0\le r_1,r_2<2^m-1$, then there are $\phi(e)/2$ choices for j. We say r_1 is equivalent to r_2, written $r_1\sim_e r_2$, if $r_2\equiv 2^k r_1 \pmod 2)^{n/e}-1$ for some k. By Proposition 6.1, a family of $(0,j)$-QF sequences is uniquely determined by the following: a choice of Galois coset of primitive elements of $GF(2^n)$; a choice of divisor e of n; a choice of j relatively prime to e; and a choice of $\sim_e$ equivalence class r, with $r\not\sim_e 1$. In addition, there is a family of $(0,j)$-QF sequences for each choice of Galois equivalence class of primitive elements, each $j<n$ relatively prime to n, and $r=1$.

The number of Galois equivalence classes of primitive elements is $\phi(2^n-1)/n$. For a given e, the number of $j<e/2$ and relatively prime to e is $\phi(e)/2$, and the number of $\sim_e$ equivalence classes is $e\phi(2^{n/e}-1)/n$, which proves the theorem.

Similar arguments apply in the second case, with e restricted to be odd. □

7 Tables

The values of $\mathbf{N}_{\rm QF}$ and the lower bound on the maximum linear span for the first few n are summarized in Table (2). The values for the linear span refer to sequences for which the correlations are bounded by $q^{(e+1)/2}+1$. This means that e is odd. If n is a power of two, then it has no odd divisors. If n is prime or is twice a prime, then r must be a power of two, so the lower bound on the linear span is $2n$. These values have been omitted from the table.

For the values included, not all families achieve the stated lower bound on the maximum linear span – this bound is simply the best that can be achieved. Also, for a given n, the maximum correlation varies as the factorization $n=me$ varies. In fact, there is a tradeoff between linear span and maximum correlation. For a fixed n, the linear span is large when e is small and m is large (and ${\rm wt}(r)$ is large). The maximum correlation, however, exceeds the square root of the period by about $2^{m/2}$ when e is odd, and by about 2^m when e is even.

n	Period	$\mathbf{N}_{\mathrm{QF}}$	Max $\lambda(\mathbf{S})$
9	511	192	$\geq$ 288
12	4 095	288	$\geq$ 2 304
15	32 767	19 800	$\geq$ 11 520
18	262 143	62 208	$\geq$ 82 944
20	1 048 575	9 600	$\geq$ 1 795
21	2 097 151	3 101 748	$\geq$ 387 072
24	16 777 215	5 840 180	$\geq$ 2 654 208
25	33 554 431	25 920 000	$\geq$ 14 670
27	134 217 727	247 947 264	$\geq$ 11 943 936
28	268 435 455	28 449 792	$\geq$ 6 591
30	1 073 741 823	17 820 000	$\geq$ 79 626 240
33	8 589 934 591	61 194 714 240	$\geq$ 350 355 456
35	34 359 738 367	56 686 248 360	$\geq$ 864 473
36	68 719 476 735	108 113 522 688	$\geq$ 2 293 235 712
39	549 755 813 887	7 717 446 434 880	$\geq$ 9 937 354 752
40	1 099 511 627 775	378 961 920 000	$\geq$ 6 386 596

Table 2: Distinct Families of $(0, j)$-QF Sequences of Period $2^n - 1$ with Minimal Correlations and Large Linear Span

m	e	wt(r)	Max. $\lambda(\mathbf{S})$	Max. Θ	No. of Families
1	36	≤ 1	≥ 72	$2^{19}+1$	$6N$
2	18	≤ 1	≥ 72	$2^{20}+1$	$3N$
3	12	≤ 2	≥ 1511	$2^{21}+1$	N
4	9	≤ 3	≥ 15974	$2^{20}+1$	N
6	6	≤ 5	≥ 415941	$2^{24}+1$	N
9	4	≤ 8	≥ 1179648	$2^{27}+1$	N
12	3	≤ 11	≥ 5314392	$2^{24}+1$	N

Table 3: Optimal Families for $n = 36$ and Various Factorizations of n

Thus the maximum correlation is small when m is small. Making a choice of factorization $n = me$ depends on which statistic is most important.

Table (3) summarizes the effects of this choice for $n = 36$. The assumption is that wt(r) is as large as possible. N denotes the number of distinct Galois cosets of primitive elements in $GF(2^{36})$. That is, $N = \phi(2^{36}-1)/36 = 725,594,112$. From the table we see that there are only three reasonable choices, $m = 1$, 4, or 12. Every other value of m gives a maximum correlation and linear span that are worse than those given by at least one of these three values.

References

[1] T. Kasami, “Weight distribution formula for some classes of cyclic codes,” Coordinated Science Laboratory, University of Illinois, Urbana, Tech. Rep. R-285 (AD632574), 1966.

[2] T. Kasami, “Weight distribution of Bose-Chaudhuri-Hocquenghem codes,” in *Combinatorial Mathematics and its Applications.* Chapel Hill, NC: University of North Carolina Press, 1969.

[3] E. L. Key, “An Analysis of the structure and complexity of nonlinear binary sequence generators,” *IEEE Trans. Info. Theory,* vol. IT-22 no. 6, pp. 732-736, Nov. 1976.

[4] A. Klapper, “*d*-form Sequences: Families of Sequences with Low Correlation Values and Large Linear Span,” University of Kentucky, Department of Computer Science Technical Report #240-93, 1993.

[5] A. Klapper, “Cross-correlations of geometric sequences in characteristic two,” *Designs, Codes, and Cryptography*, vol. 3, pp. 347-377, 1993.

[6] P. V. Kumar and R. A. Scholtz, “Bounds on the linear span of bent sequences,” *IEEE Trans. Info. Theory,* vol. IT-29, pp. 854-862, Nov. 1983.

[7] A. Lempel and M. Cohn, “Maximal families of bent sequences,” *IEEE Trans. Info. Theory,* vol. IT-28, pp. 865-868, Nov. 1982.

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

[9] J. No, *A new family of binary pseudorandom sequences having optimal periodic correlation properties and large linear span,* Doctoral Dissertation, University of Southern California, 1988.

[10] J. No and P. V. Kumar, “A new family of binary pseudorandom sequences having optimal periodic correlation properties and large linear span,” *IEEE Trans. on Info. Theory,* vol. 35, pp. 371-379, 1989.

[11] J. D. Olsen, R. A. Scholtz, and L. R. Welch, “Bent-function sequences,” *IEEE Trans. Inform. Theory,* vol. IT-28, pp. 858-864, Nov. 1982.

[12] O. Rothaus, “On bent functions,” *Journal of Combinatorial Theory Series A*, vol. 20, pp. 300-305, 1976.

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