[6] R. A. Scholtz and L. R. Welch, "Group characters: Sequences with good correlation properties," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 537–545, Sept. 1978.

[7] W. O. Alltop, "Complex sequences with low periodic correlations," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 350–354, May 1980.

[8] I. F. Blake and J. W. Mark, "A note on complex sequences with low correlations," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 814–816, Sept. 1982.

[9] M. Antweiler and L. Bömer, "Complex sequences over GF(p^M) with a two-level autocorrelation function and a large linear span," *IEEE Trans. Inform. Theory*, vol. 38, pp. 120–130, Jan. 1992.

[10] P. V. Kumar and O. Moreno, "Prime-phase sequences with periodic correlation properties better than binary sequences," *IEEE Trans.Inform. Theory*, vol. 37, pp. 603–616, May 1991.

[11] D. A. Shedd and D. V. Sarwate, "Construction of sequences with good correlation properties," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 94–97, Jan. 1979.

[12] G. H. de Visme, *Binary Sequences*. London, England: The English Universities Press, 1971.

[13] E. A. Gabidulin, "Non-binary sequences with the perfect periodic autocorrelation and with optimal periodic cross-correlation," in *ISIT'93 Symp.*, 1993, pp. 412.

[14] H. Chung and P. V. Kumar, "A new general construction for generalized bent functions," *IEEE Trans. Inform. Theory*, vol. 35, Jan. 1989.

[15] H. Miyagawa, Y. Iwatare, and H. Imai, *Coding Theory*. Tokyo, Japan: Shokodo, 1973.

[16] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications*. Rockville, MD: Computer Science Press, 1985, vol. 1.

Binary Sequences with Gold-Like Correlation but Larger Linear Span

Serdar Boztaş and P. Vijay Kumar

***Abstract*—A new construction of optimal binary sequences, identical to the well known family of Gold sequences in terms of maximum nontrivial correlation magnitude and family size, but having larger linear span is presented. The distribution of correlation values is determined. For every odd integer $r \geq 3$, the construction provides a family that contains $2^r + 1$ cyclically distinct sequences, each of period $2^r - 1$. The maximum nontrivial correlation magnitude equals $2^{(r+1)/2} + 1$. With one exception, each of the sequences in the family has linear span at least $(r^2 - r)/2$ (compared to $2r$ for Gold sequences).**

The sequences are easily implemented using a quaternary shift register followed by a simple feedforward nonlinearity.

***Index Terms*—Binary sequences, Gold sequences, sequences with low correlation, large linear span.**

I. Introduction

Code-division multiple-access (CDMA) [23] allows several users simultaneous access to a common channel by assigning a distinct *code* or *signature* sequence to each user, to enable him to distinguish his signal from those of the other users. When binary phase shift keying (BPSK) is the method of modulation employed, each symbol in the code sequence is of the form $(-1)^c$ where $c \in \{0, 1\}$. It is common to use the term code sequence to denote both the $\{\pm 1\}$ and $\{0, 1\}$ sequences. Typically, the code sequences are periodic with period L. Each user derives his distinguishing *code signal* by multiplying a common radio-frequency carrier with the $\{\pm 1\}$ code sequence. Data is mounted onto the code signal by multiplication with a $\{\pm 1\}$ *data sequence*. Transitions in the data sequence are allowed to occur only once every M code symbol durations.

In military situations where jamming is a threat, one typically chooses $M \ll L$ to prevent the threat of repeater jamming (see [23, p. 28] for instance).

The *linear span* of a periodic sequence is the length of the shortest linear-feedback shift register that can be used to generate the sequence [7], [23], [8]. As a further precaution both against jamming as well as of interception by an unfriendly receiver, the code sequences should be chosen to have large linear span l, with $l > M$. This is because, by using an algorithm such as the continued-fraction algorithm, it is possible to determine the linear recursion of a code sequence having linear span l, just from observing $2l$ consecutive bits of the sequence. This linear recursion can then be used to configure a replica of the code sequence generator for jamming and/or eavesdropping purposes. This issue is discussed in greater detail in [18, p. 863], [23, pp. 278–279 and 305], [8, p. 855], and [19].

To facilitate synchronization as well as to minimize interference due to other users, the nontrivial auto and cross correlation values of the code sequence family must be kept small. Because $M \ll L$ and since the data sequence of the incoming signal can undergo a change in sign in the midst of a correlation interval (equal to M code symbol durations) at the receiver end, both the so-called *odd* (when the data does change sign) and *even* (when the data does not change sign) partial–period (p–p) auto and cross correlations of the sequence family will affect system performance [20]. However, designing for low p–p correlation is well known to be an extremely hard problem (see [23, p. 294–295] for a discussion of previous work on the problem and [11] for some recent results). To date, there exists no design that claims to be even near-optimal with respect to the p–p correlation requirement. It is common practice to design based on full-period correlations even if the application is for a p–p situation. The resulting design is then analyzed for its p–p correlation properties. This can heuristically be justified by arguing that correlation is essentially a measure of randomness. There is some analytical support for this. For instance, the mean-square (even) p–p autocorrelation of a sequence depends only on the full-period autocorrelation of the sequence (see, for example, [13], [9], [19]).

A further property of the code sequence family that affects system vulnerability to jamming, is the relative imbalance between the 1's and 0's per period of the code sequence (see [19]). It is desirable that the sequences be as balanced in this respect as possible.

When designing for low full-period correlation, designers attempt to minimize $C_{\max}$ (the maximum, nontrivial, periodic, even correlation magnitude) for a family of given size. Lower bounds on $C_{\max}$ due to Welch [26], Sidelnikov [22], and Levenstein [12] are commonly used to judge the merits of a particular CDMA design.

The family of binary Gold sequences [3], [23], [20] are a family of $2^r + 1$ cyclically distinct sequences each of period $2^r - 1$, r odd, having $C_{\max}$ equal to $\sqrt{2^{r+1}} + 1$. By the Sidelnikov bound, the Gold family is known [20] to be optimal in terms of having

Manuscript received August 29, 1990; revised August 25, 1992. This work was supported in part by the National Science Foundation under Grant NCR-8719626 and NCR-9016077. This paper was presented in part at the 1991 IEEE International Symposium on Information Theory, Budapest, Hungary, June 1991.

S. Boztaş is with the Department of Electrical & Computer Systems Engineering, Monash University, Clayton, Victoria 3168, Australia.

P. V. Kumar is with the Communication Sciences Institute, EE-Systems, EEB 534, University of Southern California, Los Angeles, CA 90089-2565 USA.

IEEE Log Number 9400605.

the minimum possible value of $C_{\max}$ for the given family size and symbol alphabet. The maximum imbalance between 1's and 0's per period L of the code sequence is on the order of $\sqrt{L}$. This family has found widespread use as a family of signature sequences in CDMA systems (see [27], for example). However, Gold sequences have short linear span (typically $2r$ when the period equals $2^r - 1$).

A family of binary sequences, identical to the Gold family in terms of family size, correlation parameter $C_{\max}$ and range of symbol imbalance, but having typical linear span $r(r-1)/2$ is presented here. This new design arose out of earlier work by the authors (and A. R. Hammons) on families of quaternary (symbols drawn from $Z_4 := Z/4Z$) sequences [1],[2] (see also [24], [25]). This family may be viewed as being derived from a pair of elementary symmetric functions operating on r-tuples over GF(2^r) comprised of successive powers of a primitive element in the field together with its conjugates under the Galois group of GF(2^r)/GF(2).

The family of Gold sequences has the property that when one adds two distinct sequences within the family, one obtains some cyclic shift of a third member in the family. As pointed out by a referee, this property could make the Gold family more susceptible to the threat of intelligent jamming. The new design is nonlinear and hence does not share this weakness.

The connection with quaternary sequences can be exploited to provide an easy implementation of the new design using a quaternary shift register in conjunction with a simple, nonlinear, quaternary-to-binary map. By a quaternary shift-register we mean that the contents of the individual registers are now elements of Z_4 and that all recursion arithmetic is carried out modulo 4. The sequences in the new design are also strongly related to the binary, nonlinear Kerdock code (see [16] and [5]).

Sequences identical to m-sequences in terms of their autocorrelation function but having larger linear span are described in [21]. Families of sequences having the same family size and $C_{\max}$ as the small set of Kasami sequences, see [6], [23], [20], may be found in [18] and [17]. Large families of sequences with low correlation are described in [10].

The new family is defined in Section II and the correlation distribution determined. The connection with elementary symmetric functions is made at the end of this section. Section III discusses the linear span of these sequences, an example, as well as implementation of the sequence family via the link with quaternary sequences.

II. Definition and Correlation Distribution

Let $r = 2s+1$ be an odd integer ≥ 3. To simplify notation, we use $\boldsymbol{E}$ and $\boldsymbol{F}$ to denote the finite fields GF(2^r) and GF(2), respectively. Let α be a primitive element of $\boldsymbol{E}$ and let $tr(\cdot)$ denote the trace from $\boldsymbol{E}$ to $\boldsymbol{F}$.

Definition 1: The binary family $\mathcal{S}$ is defined to be the set of $M \triangleq 2^r + 1$ sequences $s_i(t)$, $0 \leq t \leq 2^r - 2$, given by

$$s_i(t) = \begin{cases} tr(v_i\alpha^t) + \sum_{l=1}^s tr((\alpha^t)^{1+2^l}) & \text{if } 1 \leq i \leq 2^r \\ tr(\alpha^t) & \text{if } i = 2^r + 1 \end{cases} \quad (1)$$

where $\{v_i/1 \leq i \leq 2^r\}$ is an enumeration of the elements of $\boldsymbol{E}$.

Clearly, each sequence $s_i(\cdot)$ in (1) has period $L \triangleq 2^r - 1$. To simplify notation, we define

$$a(t) = \sum_{l=1}^{s} tr\left((\alpha^t)^{1+2^l}\right).$$

It will be found easier to work with the related functions $p(\cdot)$ and $f_i(\cdot)$ defined over $\boldsymbol{E}$:

$$p(x) = \sum_{l=1}^{s} tr\left(x^{1+2^l}\right), \quad \forall x \in \boldsymbol{E} \quad (2)$$

and

$$f_i(x) = \begin{cases} tr(v_ix) + \sum_{l=1}^s tr(x^{1+2^l}) & \text{if } 1 \leq i \leq 2^r, \forall x \in \boldsymbol{E} \\ tr(x) & \text{if } i = 2^r + 1, \forall x \in \boldsymbol{E} \end{cases}. \quad (3)$$

The correlation values

$$\{C_{i,j}(\tau) \mid 1 \leq i,\, j \leq M,\, 0 \leq \tau \leq L-1\}$$

of the family $\mathcal{A}$ are given by

$$C_{i,j}(\tau) = \sum_{t=0}^{L-1} (-1)^{s_i(t \oplus \tau) + s_j(t)} \quad (4)$$

where $\oplus$ denotes addition modulo L. The maximum correlation magnitude is denoted by $C_{\max}$, i.e.,

$$C_{\max} = \max\{|\, C_{i,j}(\tau) \mid \text{either } i \neq j \text{ or } \tau \neq 0\}.$$

The correlation values as well as their distribution will now be determined by considering 5 cases separately.

Case A) $i = j$ and $\tau = 0$: In this trivial case,

$$C_{i,j}(\tau) = 2^r - 1.$$

Case B) $i = j = 2^r + 1$, $\tau \neq 0$: Since $s_{2^r+1}(\cdot)$ is an (see, for example, [3]) m-sequence, we have

$$C_{i,j}(\tau) = -1$$

in this case.

Case C) $i = 2^r + 1$, $j \neq 2^r + 1$ (the results apply equally of course, to the case $j = 2^r + 1$, $i \neq 2^r + 1$): In this case as well as in Case E, as an intermediate step, we will relate the relevant correlation values to the values of the Hadamard transform [4] of a Boolean function in r binary variables.

Let τ, $0 \leq \tau \leq L-1$, be fixed and note that

$$C_{2^r+1,j}(\tau) = \sum_{t=0}^{L-1} (-1)^{tr(\alpha^t[\alpha^\tau + v_j]) + a(t)}.$$

Define

$$\hat{p}(\lambda) = \sum_{x \in \boldsymbol{E}} (-1)^{p(x) + tr(x\lambda)} \, \forall \lambda \in \boldsymbol{E}. \quad (5)$$

It follows then that the listing of correlation values

$$A_\tau \triangleq \{C_{2^r+1,j}(\tau) \mid 1 \leq j \leq 2^r\}$$

is precisely the same as the list

$$\{\hat{p}(\lambda) - 1 \mid \lambda \in \boldsymbol{E}\}.$$

We emphasize that A_τ is to be regarded as a list (rather than a set) in which elements can occur with multiplicity greater than 1. We now rewrite (5) as the Hadamard transform of a Boolean function. Let $B_1 = \{\gamma_0, \gamma_1, \cdots, \gamma_{2s}\}$ be a basis for $\boldsymbol{E}$ over $\boldsymbol{F}$ and let $B_2 = \{\theta_0, \theta_1, \cdots, \theta_{2s}\}$ be a dual basis to B_1 so that,

$$tr(\gamma_i\theta_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}. \quad (6)$$

Next, define the Boolean functions $P(\cdot)$, $\hat{P}(\cdot)$ by

$$P(\underline{x}) = p\left(\sum_{i=0}^{2s} x_i\gamma_i\right),$$

and

$$\hat{P}(\underline{\lambda}) = \hat{p}\left(\sum_{i=0}^{2s} \lambda_i\theta_i\right)$$

where $x_i, \lambda_i \in \boldsymbol{F}, \underline{x} = (x_0, x_1, \cdots, x_{2s})^T$ and $\underline{\lambda} = (\lambda_0, \lambda_1, \cdots, \lambda_{2s})^T$.

Then from (5) and (6) it follows that

$$\hat{P}(\underline{\lambda}) = \sum_{\underline{x} \in \boldsymbol{F}^r} (-1)^{P(\underline{x}) + \underline{x}^T \underline{\lambda}}$$

is the desired transform equation. As a result, we have that

$$A_\tau = \{\hat{P}(\underline{\lambda}) - 1 \mid \underline{\lambda} \in \boldsymbol{F}^r\}.$$

The set

$$\{\hat{P}(\underline{\lambda}) \mid \underline{\lambda} \in \boldsymbol{F}^r\}$$

is unaltered if we replace $\hat{P}(\cdot)$ by the transform $\hat{R}(\cdot)$ of a Boolean function $R(\cdot)$ related to $P(\cdot)$ via an affine transformation of the specific form

$$R(\underline{x}) = P(A\underline{x}) + \underline{c}^T \underline{x} \tag{7}$$

in which A is an $(r \times r)$ nonsingular matrix over $\boldsymbol{F}$ and $\underline{c} \in \boldsymbol{F}^r$.

It can be shown that $P(\cdot)$ is a quadratic Boolean function (see [7] or [16, ch. 15]) and hence the set A_τ is completely determined if the rank of the associated symplectic form (See [16, ch. 15]) is known. This rank can be determined from examining the function

$$B_p(x, z) \triangleq p(x) + p(z) + p(x+z) \quad \text{for all } (x, z) \in \boldsymbol{E}^2. \tag{8}$$

We first note that B_p can be put into the form

$$B_p(x, z) = \sum_{l=1}^{s} tr(xz^{2^l} + x^{2^l} z), \tag{9}$$

which can further be simplified to

$$B_p(x, z) = tr(xz) + tr(x)tr(z). \tag{10}$$

To find the rank of the symplectic, it is sufficient to determine the number of elements $x \in \boldsymbol{E}$ such that

$$B_p(x, z) = 0 \forall z \in \boldsymbol{E},$$

i.e., the number of $x \in \boldsymbol{E}$ such that

$$tr\{z[tr(x) + x]\} = 0 \forall z \in \boldsymbol{E},$$

and since this is evidently 2, it follows that the symplectic form associated to $P(\cdot)$ has rank equal to $r - 1 = 2s$.

Now, Dickson's theorem (see [16, ch. 15]) tells us that via an affine transformation of the form in (7), $P(\cdot)$ can be put into the form

$$P(\underline{x}) = \sum_{i=1}^{s} x_i x_{i+s}.$$

The Hadamard transform values of the Boolean function $P(\cdot)$ are easily computed. As a consequence, one obtains that when $i = 2^r + 1$, as τ varies over the range $0 \le \tau \le L - 1$ and j varies over $1 \le j \le 2^r$, the resulting correlation values are distributed as follows

$$C_{i,j}(\tau) = \begin{cases} -1 & (2^r - 1)2^{2s} \text{ times.} \\ -1 + 2^{s+1}, & (2^r - 1)(2^{2s-1} + 2^{s-1}) \text{ times} \\ -1 - 2^{s+1}, & (2^r - 1)(2^{2s-1} - 2^{s-1}) \text{ times} \end{cases}. \tag{11}$$

The same distribution holds of course, for the case $j = 2^r + 1$, $1 \le i \le 2^r$ and $0 \le \tau \le L - 1$.

Case D) $\tau = 0$, $1 \le i, j \le 2^r$, $i \ne j$: In this case,

$$s_i(t \oplus \tau) + s_j(t) = tr(\alpha^t [v_i + v_j])$$

and therefore, $C_{i,j}(\tau) = -1$ always.

Case E) Finally, consider the case $\tau \ne 0, 1 \le \tau \le L - 1$ *and*$1 \le i, j \le 2^r$: Under these conditions

$$s_i(t \oplus \tau) + s_j(t) = a(t \oplus \tau) + a(t) + tr(\alpha^t [v_i \alpha^\tau + v_j]).$$

We define

$$q(x) = p(x\alpha^\tau) + p(x) \tag{12}$$

and

$$\hat{q}(\lambda) = \sum_{x \in \boldsymbol{E}} (-1)^{q(x) + tr(x\lambda)}, \ \forall \lambda \in \boldsymbol{E}.$$

It can be shown that even here, the Boolean function in r binary variables associated with $q(\cdot)$ is a quadratic form and we therefore proceed as in Case C.

Let i, τ be fixed. Then it follows that the lists

$$B_\tau = \{C_{i,j}(\tau) \mid 1 \le j \le 2^r\}$$

and

$$\{\hat{q}(\lambda) - 1 \mid \lambda \in \boldsymbol{E}\}$$

are identical. It is shown in the Appendix from an examination of

$$B_q(x, z) = q(x) + q(z) + q(x+z) \quad \text{for all } (x, z) \in \boldsymbol{E}^2, \tag{13}$$

that the symplectic form associated with $q(\cdot)$ once again has rank $r - 1 = 2s$. The distribution of the correlation values $C_{i,j}(\tau)$ can now be determined as before. It will be found then that as i, j, τ vary with $1 \le \tau \le L - 1$, $1 \le i, j \le 2^r$,

$$C_{i,j}(\tau) = \begin{cases} -1 & 2^r(2^r - 2)2^{2s} \text{ times} \\ -1 + 2^{s+1}, & 2^r(2^r - 2)(2^{2s-1} + 2^{s-1}) \text{ times} \\ -1 - 2^{s+1}, & 2^r(2^r - 2)(2^{2s-1} - 2^{s-1}) \text{ times} \end{cases}. \tag{14}$$

Collecting together the results obtained in Cases A)–E) we obtain the following.

Theorem 1: The correlation distribution for the family $\mathcal{S}$ is as follows:

$$C_{i,j}(\tau) = \begin{cases} -1 + 2^r, & 2^r + 1 \text{ times} \\ -1, & (2^{3r-1} + 2^{2r} - 2^r - 2) \text{ times} \\ -1 + 2^{s+1}, & (2^{2r} - 2)(2^{2s-1} + 2^{s-1}) \text{ times} \\ -1 - 2^{s+1}, & (2^{2r} - 2)(2^{2s-1} - 2^{s-1}) \text{ times} \end{cases}. \tag{15}$$

Corollary 2: $C_{\max} = 2^{(r+1)/2} + 1$ for the family $\mathcal{S}$ and hence the family is optimal within the class of binary sequences.

Optimality follows either from noting that the maximum correlation is identical to that in the case of Gold sequences [3], or else from an application of the Sidelnikov bound [22].

Remark: Let

$$\sigma_1(x_1, x_2, \cdots, x_r) = \sum_{i=1}^{r} x_i$$

and

$$\sigma_2(x_1, x_2, \cdots, x_r) = \sum_{j>i} x_i x_j$$

denote the first two elementary symmetric functions in the r indeterminates x_i. A little work will show that

$$tr(\alpha^t) = \sigma_1(\alpha^t, \alpha^{2t}, \cdots, \alpha^{2^{r-1t}})$$

and that

$$a(t) = \sigma_2(\alpha^t, \alpha^{2t}, \cdots, \alpha^{2^{r-1}t}).$$

Thus, one can also interpret Theorem 1 as saying that the sequences

$$\sigma_1(\alpha^t, \alpha^{2t}, \cdots, \alpha^{2^{r-1}t}) \quad \text{and} \ \sigma_2(\alpha^t, \alpha^{2t}, \cdots, \alpha^{2^{r-1}t})$$

have low cross correlation values that belong to the set

$$\{-1 \ - 1 \pm 2^{s+1}\}.$$

TABLE I

Comparison of the New Design with Some Other Binary Sequence Families. † Denotes Optimality Within the Class of Binary Sequences. ‡ Denotes that the Linear Span is as Large as the Value Given for Some Choice of Design Parameters. * Indicates that the Linear Span of all but One of the Sequences in S is at Least $r(r-1)/2$

Sequence	Period	Family Size	$C_{\max}$	Linear Span	Range of Sequence Imbalance
Gold †	$2^r - 1$ $r = 2s+1$	$2^r + 1$	$2^{(r+1)/2} + 1$	$2r$	$[1,\ 2^{(r+1)/2} + 1]$
Gold	$2^r - 1$ $r = 4s+2$	$2^r + 1$	$2^{(r+2)/2} + 1$	$2r$	$[1,\ 2^{(r+2)/2} + 1]$
Kasami† (Small Set)	$2^r - 1$ $r = 2s$	$2^{r/2}$	$2^{r/2} + 1$	$3r/2$	$[1,\ 2^{r/2} + 1]$
Kasami (Large Set)	$2^r - 1$ $r = 2s$	$\geq 2^{r/2}(2^r+1) - 1$	$2^{(r+2)/2} + 1$	$5r/2$	$[1,\ 2^{(r+2)/2} + 1]$
Bent‡	$2^r - 1$ $r = 4s$	$2^{r/2}$	$2^{r/2} + 1$	$\geq \binom{r/2}{r/4} \cdot 2^{r/4}$	1
No†‡	$2^r - 1$ $r = 2s$	$2^{r/2}$	$2^{r/2} + 1$	$\geq r \cdot 2^{r/2-2}$	$[1,\ 2^{r/2} + 1]$
New Family† S	$2^r - 1$ $r = 2s+1$	$2^r + 1$	$2^{(r+1)/2} + 1$	$r(r-1)/2*$	$[1,\ 2^{(r+1)/2} + 1]$

III. Linear Span and an Example

A. Linear Span

By a straightforward application of the results in [7] we obtain the distribution of linear span:

Theorem 3: The linear span $l_{\text{span}}(s_i(t))$ of the M sequences $s_i(t)$ in S where $1 \leq i \leq M$, has the following distribution:

$$l_{\text{span}}(s_i(t)) = \begin{cases} r(r-1)/2 & 1 \text{ time} \\ r(r+1)/2 & 2^r - 1 \text{ times} \\ r & 1 \text{ time} \end{cases}. \tag{16}$$

The linear span of r corresponds of course, to the m-sequence contained within the family.

Table I compares the family S with some other well known binary families. (This list is not claimed to be exhaustive). Note that in terms of family size, $C_{\max}$ and symbol imbalance, the new design is identical to the family of Gold (r odd) sequences. The new design of course, has larger linear span.

B. An example of S

Let $r = 5$, and let the Galois field GF(32) be generated by the primitive element α that satisfies

$$\alpha^5 + \alpha^2 + 1 = 0.$$

Then the sequences $s_i(t)$ in S are given by

$$s_i(t) = \begin{cases} tr(v_i\alpha^t) + tr(\alpha^{3t}) + tr(\alpha^{5t}) & \text{if } 1 \leq i \leq 32 \\ tr(\alpha^t) & \text{if } i = 33 \end{cases}$$

where the $\{v_i\}$ vary over all of GF(32) as i ranges between 1 and 32. Using the finite-field table given in [16, p. 10], for example, one finds that as t varies over $0 \leq t \leq 2^r - 2$,

$$\{tr(\alpha^t)\} = 10010\ 11001\ 11110\ 00110\ 11101\ 01000\ 0 \tag{17}$$

$$\{tr(\alpha^{3t}) + tr(\alpha^{5t})\} = 00010\ 01100\ 01111\ 10101\ 01111\ 11111\ 1. \tag{18}$$

The top sequence is $s_{33}(t)$. By adding various cyclic shifts of the top sequence to the bottom sequence one obtains 31 additional sequences $s_i(t)$. The bottom sequence then by itself, completes the family.

The correlation distribution of this example family was experimentally verified to conform with Theorem 1 above. This distribution is also given below

$$C_{i,j}(\tau) = \begin{cases} 31, & 33 \text{ times} \\ -1, & 17374 \text{ times} \\ +7, & 10220 \text{ times} \\ -9, & 6132 \text{ times} \end{cases}. \tag{19}$$

C. A Quaternary Implementation

The sequences in the new design can be simply implemented using a quaternary shift-register in conjunction with a simple, nonlinear, quaternary to binary mapping.

Let α be a primitive element of GF(2^r) as before. Let $f(x) = \sum_{i=0}^{r} f_i x^i$ be the minimum polynomial of α over GF(2).

Let $F(x) := \sum_{i=0}^{r} F_i x^i \in Z_4[x]$ with $f_i = F_i$ (when both are regarded as integers in Z). Let $F_e(x)$ and $F_o(x)$ be the polynomials corresponding to the even and odd exponents of $F(x)$, respectively. Then $F(x) = F_e(x) + F_o(x)$. Let the monic polynomial $G(x) \in Z_4[x]$ be defined by

$$G(x^2) = \pm([F_e(x)]^2 - [F_o(x)]^2).$$

Then it is known (see [5]) that $G(x)$ is the unique polynomial in $Z_4[x]$ satisfying the following:

- $G(x)$ is irreducible,
- the smallest integer e for which $G(x)$ divides $x^e - 1$ is $e = 2^r - 1$ and
- $G(x) = f(x)$ mod 2.

(This makes $G(x)$ a primitive, basic irreducible in $Z_4[x]$ in the terminology of [15] and [5] and this particular derivation is called Graeffe's method). Let $\{u(t) \mid u(t) \in Z_4\}$ denote the quaternary sequence satisfying the linear recurrence whose characteristic polynomial is $G(x)$, i.e., if

$$G(x) = x^r + \sum_{i=0}^{r-1} G_i x^i,$$

then

$$u(t+r) + \sum_{i=0}^{r-1} G_i u(t+i) = 0, \quad \forall t \geq 0.$$

Further $\forall t \geq 0$, let

$$u(t) = u_1(t) + 2u_2(t), \quad u_i(t) \in \{0, 1\},\ i = 1, 2,$$

be the 2-adic expansion of $u(t)$. We will refer to the component $u_2(t)$ in this expansion as the most significant bit (MSB) component.

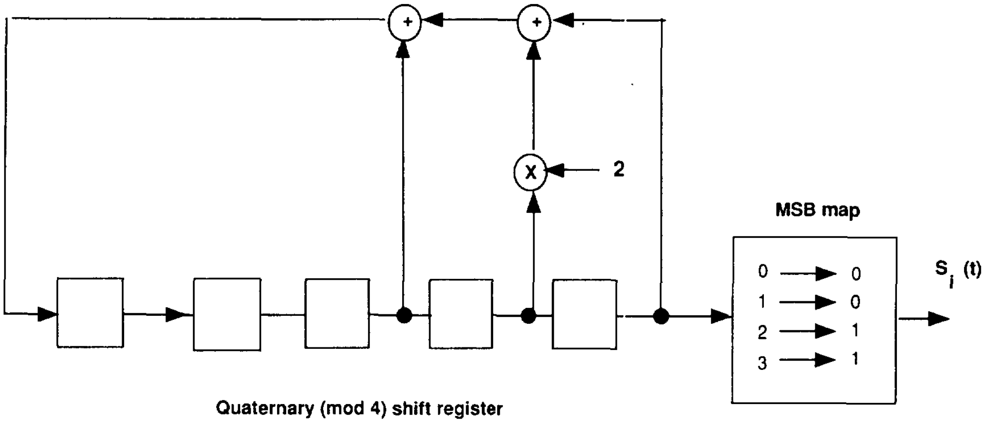

Fig. 1. A quaternary implementation of the new design $\mathcal{S}$ for $r = 5$. Different sequences within the family can be generated simply by changing the initial contents of the shift register.

The output of the quaternary shift-register naturally depends upon the initial contents of the shift register. As it turns out (see [5] or [2]), this output is always a quaternary periodic sequence having period $2^r - 1$ (except of course for the case when the initial contents of the shift register are all zero). Since there are $4^r - 1$ possible distinct nonzero initializations of the shift register, it follows that the shift register generates $2^r + 1$ sequences which are pairwise, cyclically distinct (i.e., one sequence is not a cyclic shift of the second). In fact, as shown below, the binary MSB components of these $2^r + 1$ quaternary sequences, correspond precisely, to the $2^r + 1$ binary sequences in the new design $\mathcal{S}$.

Let J denote the subset of Z_4^r of size 2^r consisting of all the r-tuples in Z_4^r, all of whose components are either 0 or 1, i.e.,

$$J = \{(z_0, z_1, \cdots, z_{r-1}) \mid z_i \in \{0, 1\} \subset Z_4\}.$$

Let $\underline{0}$ denote all-zero r-tuple over Z_4. Let $\underline{x} = (x_0, x_1, \cdots, x_{r-1})$ denote the initial contents of the quaternary shift register. Then $\underline{x}$ has the unique decomposition

$$\underline{x} = \underline{x}_1 + 2\underline{x}_2, \quad \underline{x}_1, \underline{x}_2 \in J.$$

It is shown in [2] (see also [5]) that the 2-adic components $u_1(t)$, $u_2(t)$ of the output $u(t)$ of the quaternary shift register are always of the form

$$u_1(t) = tr(\gamma\alpha^t) \tag{20}$$

$$u_2(t) = p(\gamma\alpha^t) + tr(\eta\alpha^t) \tag{21}$$

where $p(\cdot)$ is as defined in (2) and where the constants γ, η lying in GF(2^r) depend upon the initial contents $\underline{x}$ of the shift register in the following way:

- (both $\gamma = 0$ and $\eta = 0$) iff $\underline{x} = \underline{0}$,
- $\gamma = 0$ iff $\underline{x}_1 = 0$,
- for fixed $\underline{x}_1$, as $\underline{x}_2$ varies over all of J, η varies over all of GF(2^r).

From this it is clear how to generate the sequences in the new design as follows:

1) given length $2^r - 1$, pick a primitive element α in GF(2^r);
2) let $f(x)$ be the minimum polynomial of α and "lift" $f(x)$ to a basic irreducible $G(x) \in Z_4[x]$ as described above,
3) set up a quaternary shift register having characteristic polynomial $G(x)$
4) to generate s_{2^r+1}, choose initial condition $\underline{x} \neq \underline{0}$, having 2-adic components $\underline{x}_1$, $\underline{x}_2$, with $\underline{x}_1 = \underline{0}$ and $\underline{x}_2 \neq \underline{0}$;
5) to generate $s_i(t)$, $1 \leq i \leq 2^r$, choose $\underline{x}$ with $\underline{x}_1 \neq \underline{0}$ and vary $\underline{x}_2$ over all of J.

As an example, Fig. 1 shows the circuitry needed to generate $\mathcal{S}$ for the case when $r = 5$ and α is a primitive element with minimum polynomial

$$f(x) = x^5 + x^2 + 1.$$

Then Graeffe s method applied to $f(x)$ yields

$$G(x) = x^5 + 3x^2 + 2x + 3.$$

This leads to the quaternary linear recursion

$$u(t + 5) = u(t + 2) + 2u(t + 1) + u(t).$$

To generate the binary m-sequence $s_{33}(t) \in \mathcal{S}$ one can choose for example, the initial condition

$$\underline{x} = (0, 0, 0, 0, 2)$$

which yields the binary MSB component sequence

$$\{s_{33}(t + \tau')\} = 00001\ 00101\ 10011\ 11100\ 01101\ 11010\ 1$$

for some τ' as t varies over $0 \leq t \leq 30$.

Setting $\underline{x} = (0, 0, 0, 0, 1)$, for example, generates

$$\{s_i(t + \tau'')\} = 00000\ 00010\ 00101\ 10111\ 00011\ 10011\ 0$$

for some i, $1 \leq i \leq 2^r$. These results can of course be verified to be consistent with the results in Section III-B.

Appendix

A. Rank of the Symplectic form associated with $q(\cdot)$

We have

$$B_q(x, z) = q(x) + q(z) + q(x + z), \tag{22}$$

or

$$B_q(x, z) = p(x) + p(z) + p(x + z) + p(xy) + p(zy) + p((x + z)y) \tag{23}$$

by noting from (12) that

$$q(x) = p(x) + p(xy)$$

with $y \triangleq \alpha^r$. Applying (10), we obtain

$$B_q(x,\, z) = tr(y^2xz) + tr(yx)tr(yz) + tr(xz) + tr(x)tr(z). \quad (24)$$

To find the rank of the quadratic form $q(x)$ it is sufficient to find the *number* of elements $x \in \boldsymbol{E}$ such that

$$B_q(x,\, z) = 0 \quad \text{for all } z \in \boldsymbol{E}. \quad (25)$$

Now, x is one such element if and only if

$$tr\{z[tr(x) + ytr(xy) + x + xy^2]\} = 0 \quad \text{for all } z \in \boldsymbol{E}. \quad (26)$$

This happens if and only if

$$x + xy^2 + tr(x) + ytr(xy) = 0. \quad (27)$$

To determine the number of $x \in \boldsymbol{E}$ that satisfy (27), we will need a simple result on finite fields. Recall that $y = \alpha^r$ and $r \neq 0$ implies that $y \neq 1$, $y \neq 0$. Given any element $\theta \in \boldsymbol{E}$, $\theta \neq 1$, we have

$$tr\left(\frac{\theta}{1+\theta^2}\right) = tr\left(\frac{1}{1+\theta^2}\right) + tr\left(\frac{1+\theta}{1+\theta^2}\right) = 0. \quad (28)$$

There are four cases to be considered:

a) $tr(x) = tr(yx) = 0$: In this case, (27) becomes $x + xy^2 = 0$ or $x(1+y^2) = 0$, which implies that $x = 0$, since $y \notin \boldsymbol{F}$ implies that $(1+y^2) \neq 0$.

b) $tr(x) = 1$, $tr(xy) = 0$: We now have $x + xy^2 + 1 = 0$, or $x(1+y^2) = 1$, or $x = (1/(1+y^2))$. This value of x is a valid solution to (27) iff $tr(1/(1+y)) = 1$. Note that by (28), the condition $tr(xy) = 0$ provides no additional information.

c) $tr(x) = 0$, $tr(xy) = 1$: We now have $x + xy^2 + y = 0$, or $x = (y/(1+y^2))$. Similar to the previous case, this value of x is a valid solution to (27) iff $tr(1/(1+y)) = 0$. Hence the cases b) and c) are mutually exclusive and depending on the value of y, only one of them can occur.

d) $tr(x) = 1$, $tr(xy) = 1$: We now have $x + xy^2 + 1 + y = 0$, or $x(1+y^2) = 1+y$, or $x = (1/(1+y))$. This is possible only if $tr(1/(1+y)) = 1$ and $tr(y/(1+y)) = 1$. However, these two equations are contradictory since r is odd. Hence this case can never occur.

Therefore, for every value of y, $y \neq 0$, there are precisely two solutions to (27) and therefore, the rank of the symplectic associated to $q(\cdot)$ is $2s = r - 1$.

References

[1] S. Boztaş, "Near-optimal four-phase sequences and optimal binary sequences for CDMA," Ph.D. dissertation, Univ. Southern California, May 1990.

[2] S. Boztaş, A. R. Hammons, and P. V. Kumar, "4-phase sequences with near-optimum correlation properties," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1101–1113, May 1992.

[3] R. Gold, "Optimal binary sequences for spread spectrum multiplexing," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 154–156, Oct. 1967.

[4] S. W. Golomb, *Shift Register Sequences.* San Francisco, CA: Holden-Day, 1967; revised edition: Aegean Park Press, Laguna Hills, CA 1982.

[5] A. R. Hammons, Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé, "The Z_4-linearity of the Kerdock, Preparata, Goethals and related codes," *IEEE Trans. Inform. Theory*, see this issue, pp. xxx–xxx.

[6] T. Kasami, "Weight distribution formula for some class of cyclic codes," Coordinated Sci. Lab., Univ. Illinois, Urbana, IL,Tech. Rep. R-285 (AD 632574), 1966.

[7] E. L. Key, "An analysis of the structure and complexity of nonlinear binary sequence generators," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 732–736, Nov. 1976.

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

[9] P. V. Kumar, "The partial-period correlation moments of arbitrary, binary sequences," in *1985 IEEE Global Telecommun. Conf. Rec.*, New Orleans, LA, December 2–5, 1985, pp. 499–503.

[10] P. V. Kumar, T. Helleseth, A. R. Calderbank, and A. R. Hammons, Jr., "Large four-phase sequence families with low correlation for CDMA," preprint, 1992.

[11] P. V. Kumar and V. K. Wei, "Minimum distance of logarithmic and fractional partial m-sequences," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1474–1482, Sept. 1992.

[12] V. I. Levenstein, "Bounds on the cardinality of a code with bounded modulus of the inner product," *Sov. Math. Dokl.*, vol. 25, no. 2, pp. 526–531, 1982.

[13] J. H. Lindholm, "An analysis of the pseudo-randomness properties of subsequences of long m-sequences," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 569–576, 1968.

[14] C. M. Liu and P. V. Kumar, "On lower bounds to the maximum correlation of complex roots-of-unity sequences," *IEEE Trans. Inform. Theory*, vol. 36, pp. 633–640, May 1990.

[15] B. R. MacDonald, *Finite Rings with Identity.* New York: Marcel-Dekker, 1974.

[16] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* New York: North-Holland, 1977.

[17] J. S. No and P. V. Kumar, "A new family of binary pseudorandom sequences having optimal correlation properties and large linear span," *IEEE Trans. Inform. Theory*, vol. 35, pp. 371–379, Mar. 1989.

[18] J. D. Olsen, R. A. Scholtz, and L. R. Welsch, "Bent-function sequences," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 858–864, Nov. 1982.

[19] M. P. Ristenbatt and J. L. Daws, Jr., "Performance criteria for spread-spectrum communications," *IEEE Trans. Commun.*, vol. COM-25, pp. 756–763, Aug. 1977.

[20] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, vol. 68, pp. 593–620, May 1980.

[21] R. A. Scholtz and L. R. Welch, GMW sequences, *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 548–553, May 1984.

[22] V. M. Sidelnikov, "On mutual correlation of sequences," *Sov. Math. Doklady*, vol. 12, pp. 197–201, 1971.

[23] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread-Spectrum Communications, Volume 1.* Rockville, MD: Computer Science, 1985.

[24] P. Solé, 'A quaternary cyclic code, and a family of quadriphase sequences with low correlation properties," *Coding Theory and Applications, Lecture Notes in Computer Science.* New York: Springer-Verlag, 1989, pp. 193–201.

[25] P. Udaya and M. U. Siddiqi, "Large linear complexity sequences over Z_4 for quadriphase modulated communication systems having good correlation properties," in *Proc. IEEE Int. Symp. Inform. Theory*, Budapest, Hungary, June 24–28, 1991, p. 386.

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

[27] R. Ziemer and R. Peterson, *Digital Communication and Spread-Spectrum Communication Systems.* New York: McMillan, 1985.