A New Binary Sequence Family with Low Correlation and Large Size

Nam Yul Yu and Guang Gong, *Member, IEEE*

***Abstract*— For odd $n = 2l+1$ and an integer ρ with $1 \le \rho \le l$, a new family $\mathcal{S}_o(\rho)$ of binary sequences of period $2^n - 1$ is constructed. For a given ρ, $\mathcal{S}_o(\rho)$ has maximum correlation $1 + 2^{\frac{n+2\rho-1}{2}}$, family size $2^{n\rho}$, and maximum linear span $\frac{n(n+1)}{2}$. Similarly, a new family of $\mathcal{S}_e(\rho)$ of binary sequences of period $2^n - 1$ is also presented for even $n = 2l$ and an integer ρ with $1 \le \rho < l$, where maximum correlation, family size, and maximum linear span are $1+2^{\frac{n}{2}+\rho}$, $2^{n\rho}$, and $\frac{n(n+1)}{2}$, respectively. The new family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) contains Boztas and Kumar's construction [1] (or Udaya's [19]) as a subset if m-sequences are excluded from both constructions. As a good candidate with low correlation and large family size, the family $\mathcal{S}_o(2)$ is discussed in detail by analyzing its distribution of correlation values.**

***Index Terms*— Family of binary sequences, large family size, linear span, sequences with low correlation.**

I. Introduction

In code division multiple access (CDMA) communication systems, pseudo-noise *sequences* are assigned to distinct users in a common channel at the same time [18]. To distinguish each user and minimize mutual interference, we must have low *crosscorrelation* between distinct sequences. Furthermore, we must also have low *autocorrelation* between a sequence and its time shifted version in order to acquire the accurate phase information at the receiver. The capacity of the CDMA system can be increased by obtaining an increased number of sequences which support a larger number of distinct users. Consequently, a family of sequences with low correlation and large family size plays important roles in CDMA communication systems.

The construction of a family of binary sequences of period $2^n - 1$ is based on the combination of a binary m-sequence and its decimations such that the resulting sequences have low correlation achieving the well known lower bounds derived by Welch [20], Sidelnikov [17], and Levenshtein [10]. For odd n, Gold sequences [3] constitute one of the families with optimal correlation achieving the Sidelnikov bound. A family of Kasami (small set) sequences [6] gives sequences with optimal correlation achieving the Welch's lower bound for even n.

In order to obtain binary sequences with large linear span as well as low correlation, Boztas and Kumar constructed a new family of binary sequences for odd n, so called *Gold-like sequences* [1]. It has the same period, family size, and maximum correlation as those of the family of Gold sequences, but larger linear span giving better potential cryptographic property. Similarly, Udaya constructed a new family of binary sequences with low correlation but large linear span for even n [19]. In [7], Kim and No generalized these two constructions at the price of the decrease of maximum linear span and the increase of maximum correlation.

The other known approaches are summarized as follows. In [2], Chang *et al.* showed that a binary cyclic code based on three-term sequences [13] has five-valued nonzero weight distribution, which is identical to the dual code of the triple error correcting BCH code. From these cyclic codes, equivalent families of binary sequences with six-valued correlation of maximum $1 + 2^{\frac{n+3}{2}}$, family size 2^{2n}, and maximum linear span $3n$, can be constructed. In [12], it is shown that Z_4-linear binary codes become nonlinear cyclic codes after a proper permutation. From these codes, Shanbhag, Kumar, and Helleseth presented a new generalized construction of binary sequence families in [16], including Kerdock and Delsarte-Goethals sequences in [5].

In this paper, a new family $\mathcal{S}_o(\rho)$ of binary sequences of period $2^n - 1$ is constructed for odd $n = 2l+1$ and an integer ρ with $1 \le \rho \le l$. For a given ρ, maximum correlation of sequences in $\mathcal{S}_o(\rho)$ is $1 + 2^{\frac{n+2\rho-1}{2}}$ and its family size is $2^{n\rho}$. Similarly, a new family $\mathcal{S}_e(\rho)$ of binary sequences of period $2^n - 1$ is also presented for even $n = 2l$ and an integer ρ with $1 \le \rho < l$, where maximum correlation and family size are $1 + 2^{\frac{n}{2}+\rho}$ and $2^{n\rho}$, respectively. The maximum and minimum linear spans of sequences in both $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ are $\frac{n(n+1)}{2}$ and $\frac{n(n-2\rho+1)}{2}$, respectively. As ρ increases, we obtain a new family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) of exponentially increased family size with maximum correlation linearly increased from its optimal value. Since the family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) contains $\mathcal{S}_o(\rho-1)$ (or $\mathcal{S}_e(\rho-1)$) as a subset, it contains $\mathcal{S}_o(1)$ (or $\mathcal{S}_e(1)$) as a subset. Here, $\mathcal{S}_o(1)$ is the family of Gold-like sequences constructed by Boztas and Kumar, and $\mathcal{S}_e(1)$ is the one constructed by Udaya, where m-sequences are excluded in both constructions.

For a specific application, we can choose a proper value ρ and the corresponding family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$). For example, a small value of ρ can be chosen if low correlation is more crucial than large family size in the application. If large family size is more important, on the other hand, we can choose a large value of ρ. The flexibility due to ρ allows our new sequence family to have adaptive family size and maximum correlation for practical applications. Furthermore, our new sequence family has good potential cryptographic property with large linear span. The implementation of $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) is extremely easy by summing linear feedback shift

This work was supported by NSERC Grant RGPIN 227700-00.

The authors are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (email: nyyu@engmail.uwaterloo.ca, ggong@calliope.uwaterloo.ca).

register (LFSR) outputs. The family $\mathcal{S}_o(2)$ with maximum correlation $1+2^{\frac{n+3}{2}}$ and family size 2^{2n} is a good example for compromise between correlation and family size.

This paper is organized as follows. In Section II, we give some preliminaries on concepts and definitions of codes and sequences. Also, we review a weight distribution of a linear cyclic subcode of the second order *Reed-Muller* code [11], which will be used to investigate a correlation distribution of sequences in our new family. In Section III, we present new families $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ of binary sequences of period 2^n-1 for odd and even n, respectively, and analyze correlation and linear span of each family. In Section IV, the distribution of correlation values of sequences in $\mathcal{S}_o(2)$ is derived in terms of the weight distribution of a linear cyclic subcode of the second order Reed-Muller code. In Section V, an example of sequences in $\mathcal{S}_o(2)$ is given and implementation based on LFSRs is discussed. Concluding remarks and some observations are given in Section VI.

II. Preliminaries

The following notations will be used throughout this paper.

- $\mathbb{F}_q = GF(q)$ is the finite field with q elements and $\mathbb{F}_q^*$, the multiplicative group of $\mathbb{F}_q$.
- $\mathbb{F}_2^n$ is a vector space over $\mathbb{F}_2 = \{0,1\}$ with a set of all binary n-tuples.
- Let n, m be positive integers and $m|n$. The trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$ is denoted by $Tr_m^n(x)$, i.e.,

$$Tr_m^n(x) = x + x^{2^m} + \cdots + x^{2^{m(\frac{n}{m}-1)}}, \ x \in \mathbb{F}_{2^n}.$$

$Tr_1^n(x)$ is simply denoted as $Tr(x)$ if the context is clear.

A. Basic concepts

(a) Boolean function

Let $\mathbf{x} = (x_1, \cdots, x_n)$ be a vector in $\mathbb{F}_2^n$ with $x_i \in \mathbb{F}_2$ and $f(\mathbf{x})$ a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. Then, the function $f(\mathbf{x})$ taking on values 0 or 1 is called a *Boolean function* [11]. A Boolean function consists of a sum of all possible products of x_{i_j}'s with coefficients 0 or 1 [4], i.e.,

$$f(\mathbf{x}) = f(x_1, \cdots, x_n) = \sum_{1 \le j \le n} c_{i_1 i_2 \cdots i_j} x_{i_1} x_{i_2} \cdots x_{i_j}, \quad c_{i_1 i_2 \cdots i_j} \in \mathbb{F}_2 \tag{1}$$

where maximum value of j with nonzero $c_{i_1 i_2 \cdots i_j}$ is called the *degree* of the Boolean function $f(\mathbf{x})$. (1) is called an *algebraic normal form* of a Boolean function [4].

(b) Reed-Muller codes

For $0 \le r \le n$, the *rth order Reed-Muller (RM) code* $R(r,n)$ of length $N = 2^n$ is defined by a set of all vectors $F(\mathbf{y}_1, \cdots, \mathbf{y}_n)$ of length N given by [11]

$$F(\mathbf{y}_1, \cdots, \mathbf{y}_n) = c_0 \underline{\mathbf{1}} + \sum_{1 \le j \le r} c_{i_1 i_2 \cdots i_j} \mathbf{y}_{i_1} \cdot \mathbf{y}_{i_2} \cdots \mathbf{y}_{i_j}, \quad c_0, c_{i_1 i_2 \cdots i_j} \in \mathbb{F}_2 \tag{2}$$

where $\underline{\mathbf{1}} = (1, \cdots, 1)$ of length N, $\mathbf{y}_1, \cdots, \mathbf{y}_n$ are basis vectors of length N for $R(1,n)$, $\{i_1, \cdots, i_j\} \subset \{1, 2, \cdots, n\}$, and $\mathbf{y}_i \cdot \mathbf{y}_j$ is the vector whose kth element is the product of the kth elements of $\mathbf{y}_i$ and $\mathbf{y}_j$, respectively. Indeed, $R(0,n)$ is all zero or all one vector of length N, and $R(1,n)$ is always the dual of the extended Hamming code which is also obtained from a Sylvester-type Hadamard matrix [11]. In (2), the kth element of a codeword in $R(r,n)$ is given by a Boolean function $f(y_{1,k}, \cdots, y_{n,k})$ with degree of at most r, where $y_{1,k}, \cdots, y_{n,k}$ are the kth elements of $\mathbf{y}_1, \cdots, \mathbf{y}_n$, respectively. If we remove the kth components corresponding to $y_{1,k} = \cdots = y_{n,k} = 0$, then we obtain a *punctured* RM code $R(r,n)^*$ for $0 \le r \le n-1$, where each codeword has length $2^n - 1$.

(c) Trace representation of a binary periodic sequence

Let $\mathcal{S}$ be a set of all binary sequences of period $v | (2^n - 1)$ and $\mathcal{F}$ be a set of all functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. For any function $f(x) \in \mathcal{F}$, $f(x)$ can be represented as [4]

$$f(x) = \sum_{i=1}^{r} Tr_1^{n_i}(A_i x^{v_i}), \quad A_i \in \mathbb{F}_{2^{n_i}}$$

where v_i is a coset leader of a cyclotomic coset modulo $2^{n_i}-1$, and $n_i | n$ is the size of the cyclotomic coset containing v_i. For any sequence $\mathbf{a} = \{a_i\} \in \mathcal{S}$, there exists $f(x) \in \mathcal{F}$ such that

$$a_i = f(\alpha^i), \quad i = 0, 1, \cdots$$

where α is a primitive element of $\mathbb{F}_{2^n}$. Then, $f(x)$ is called a *trace representation* of $\mathbf{a}$. The linear span of the sequence $\mathbf{a}$ is equal to $\sum_{i, A_i \neq 0} n_i$, or equivalently the degree of the shortest linear feedback shift registers that can generate $\mathbf{a}$.

(d) Weight and exponential sum of a binary codeword or sequence

Let $\mathbf{a} = (a_0, a_1, \cdots, a_{v-1})$ be a binary codeword of length v, or a binary sequence of period v. The number of 1's in the codeword is called a *(Hamming) weight*. Clearly, the total sum of additive characters $(-1)^{a_i}$ of a codeword or sequence $\mathbf{a}$ with weight w is given by

$$\sum_{i=0}^{v-1} (-1)^{a_i} = v - 2w.$$

If $v = 2^n - 1$ and $\mathbf{a}$ is represented by a trace representation $f(x)$, i.e., $a_i = f(\alpha^i)$ for a primitive element α of $\mathbb{F}_{2^n}$, then the *exponential sum of* $f(x)$ *over* $\mathbb{F}_{2^n}$ is given by

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)} = 1 + \sum_{i=0}^{v-1} (-1)^{a_i} = 2^n - 2w$$

with $f(0) = 0$. Hence, the exponential sum of $f(x)$ has one-to-one correspondence with the weight of a codeword or sequence given by $f(x)$.

(e) Correlation of binary sequences

Let $\mathbf{a}$ and $\mathbf{b}$ be binary sequences of period v. The correlation of $\mathbf{a}$ and $\mathbf{b}$ is defined by

$$C_{\mathbf{a},\mathbf{b}}(\tau) = \sum_{i=0}^{v-1} (-1)^{a_i + b_{i+\tau}}, \quad 0 \le \tau \le v-1$$

where τ is a phase shift of the sequence $\mathbf{b}$ and the indices are reduced modulo v. If $\mathbf{b}$ is cyclically equivalent to $\mathbf{a}$, i.e.,

$\mathbf{b} = (a_i, a_{i+1}, \cdots, a_{i-1})$, $C_{\mathbf{a},\mathbf{b}}(\tau)$ is the *autocorrelation* of $\mathbf{a}$, $C_{\mathbf{a}}(\tau)$ for short. Otherwise, $C_{\mathbf{a},\mathbf{b}}(\tau)$ is the *crosscorrelation* of $\mathbf{a}$ and $\mathbf{b}$. We may write $a_i + b_{i+\tau} = f(\alpha^i)$ for $f(x)$ with a given τ and a primitive element α of $\mathbb{F}_{2^n}$. Then,

$$C_{\mathbf{a},\mathbf{b}}(\tau) = \sum_{i=0}^{v-1} (-1)^{a_i + b_{i+\tau}} = -1 + \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)}$$

with $f(0) = 0$. Hence, $C_{\mathbf{a},\mathbf{b}}(\tau)$ can be presented by the exponential sum of $f(x)$.

(f) Binary signal set

For r binary cyclically distinct sequences of period v, i.e., $\mathbf{s}^{(j)} = (s_0^{(j)}, \cdots, s_{v-1}^{(j)})$ with $0 \le j < r$, let $\mathcal{S} = \{\mathbf{s}^{(0)}, \cdots, \mathbf{s}^{(r-1)}\}$ and

$$C_{\max} = \max |C_{\mathbf{s}^{(i)},\mathbf{s}^{(j)}}(\tau)| \text{ for any } 0 \le \tau < v,\ 0 \le i, j < r$$

where $\tau \neq 0$ if $i = j$. Clearly, $C_{\max}$ is maximum of all nontrivial auto- and crosscorrelations of the sequences in $\mathcal{S}$. The set $\mathcal{S}$ will be called a $(v, r, C_{\max})$ *signal set* or *family of sequences*, where r is the *set size* or *family size*, and $C_{\max}$ is the *maximum correlation magnitude* of $\mathcal{S}$.

B. Weight distribution of a linear subcode of $R(2,n)^*$

In this paper, we consider a componentwise sum of a pair of binary sequences, where the sum is equivalent to a codeword of a linear cyclic subcode of the punctured second order Reed-Muller code. Thus, we can apply the weight distribution of the subcode for the distribution of correlation values of the sequences.

(a) Weight distribution of a codeword set with rank $2h$

For odd $n = 2l + 1$, we consider a codeword given by

$$f(x) = Tr(\eta_0 x) + \sum_{j=1}^{l} Tr(\eta_j x^{1+2^j}), \quad x \in \mathbb{F}_{2^n}^* \tag{3}$$

where each η_j with $0 \le j \le l$ is an element in $\mathbb{F}_{2^n}$. For even $n = 2l$, on the other hand, we consider a codeword given by

$$f(x) = Tr(\eta_0 x) + \sum_{j=1}^{l-1} Tr(\eta_j x^{1+2^j}) + Tr_1^l(\eta_l x^{1+2^l}), \quad x \in \mathbb{F}_{2^n}^* \tag{4}$$

where each $\eta_j \in \mathbb{F}_{2^n}$ for $0 \le j \le l-1$, and $\eta_l \in \mathbb{F}_{2^l}$. With respect to a basis $\{\beta_1, \cdots, \beta_n\}$ of $\mathbb{F}_{2^n}$, $x = \sum_{i=1}^n x_i \beta_i$ is an expansion of x with $x_i \in \mathbb{F}_2$ for all i. Applying this expansion to (3) or (4), we see that $f(x) = f(\sum_{i=1}^n x_i \beta_i)$ is equivalent to a Boolean function of degree less than or equal to 2, and it may be written as follows:

$$f(\mathbf{x}) = f(x_1, \cdots, x_n) = \mathbf{x}\mathbf{Q}\mathbf{x}^T + \mathbf{w}\cdot\mathbf{x}^T, \quad \mathbf{x} \in \mathbb{F}_2^n \backslash \{\underline{\mathbf{0}}\} \tag{5}$$

where $\mathbf{Q}$ is an $n \times n$ binary upper triangular matrix, $\mathbf{w}$ is a binary vector in $\mathbb{F}_2^n$, and $\underline{\mathbf{0}} = (0, \cdots, 0)$ of length n. Obviously, $\mathbf{Q}$ is determined by η_j's for nonzero j, and $\mathbf{w}$ by η_0. While $\mathbf{x} = (x_1, \cdots, x_n)$ runs through each nonzero binary n-tuple in $\mathbb{F}_2^n$, $f(\mathbf{x})$ produces each element of a codeword of length $2^n - 1$ in $R(2,n)^*$. Equivalently, $f(x)$ forms a codeword in $R(2,n)^*$ for $x \in \mathbb{F}_{2^n}^*$.

For a given nonzero $\mathbf{Q}$, it is well known that the weight distribution of a set of codewords of a *quadratic Boolean function* $f(\mathbf{x})$ for all $\mathbf{w}$ is determined by a rank of a *symplectic matrix* $\mathbf{B} = \mathbf{Q} + \mathbf{Q}^T$ [11]. Equivalently, we can consider a *symplectic form* $B_f(x,z) = f(x) + f(x+z) + f(z)$ associated with $f(x)$ for given η_j's with $1 \le j \le l$ [5]. We list the following fact regarding the distribution of exponential sums of $f(x)$.

Fact 1 (Theorem 6.2 in [5]): Let η_j's of $f(x)$ in (3) or (4) be given such that at least one η_j is nonzero for $1 \le j \le l$, where $l = \lfloor \frac{n}{2} \rfloor$. For an integer h with $1 \le h \le l$, if $B_f(x,z)$ has a rank $2h$, or equivalently $B_f(x,z) = 0$ has 2^{n-2h} solutions in $x \in \mathbb{F}_{2^n}$ for all $z \in \mathbb{F}_{2^n}^*$, then the exponential sum of $f(x)$ takes on values of 0 and $\pm 2^{n-h}$ for all $\eta_0 \in \mathbb{F}_{2^n}$, and its distribution is given by

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)} = \begin{cases} 0, & 2^n - 2^{2h} \text{ times} \\ +2^{n-h}, & 2^{2h-1} + 2^{h-1} \text{ times} \\ -2^{n-h}, & 2^{2h-1} - 2^{h-1} \text{ times} \end{cases}$$

with $f(0) = 0$.

(b) Weight distribution of a linear subcode with multiple ranks

For a set of distinct nonzero $\mathbf{Q}$'s, we can consider a set of codewords given by $f(\mathbf{x})$ in (5) for all $\mathbf{w}$. Equivalently, we can consider a set of codewords given by $f(x)$ for distinct sets of η_j's such that at least one η_j is nonzero in each set of η_j's for $1 \le j \le l$. Then, $f(x)$ may have distinct multiple ranks each of which corresponds to each set of given η_j's. Consequently, the exponential sum of $f(x)$ can take on values of 0 and $\pm 2^{n-h}$ for each possible h from Fact 1. If $f(x)$ further constitutes a linear subcode for the sets of η_j's, we may use the weight distribution of the subcode in order to investigate the distribution of exponential sums of $f(x)$.

Next, we specify the known weight distributions of two linear cyclic subcodes of $R(2,n)^*$ for odd n, which will be used in a later section for determining the correlation distribution of a new family of sequences. For a, b in $\mathbb{F}_{2^n}$ and k with $\gcd(k,n) = 1$ for odd n, a linear cyclic subcode C_G given by $f(x) = Tr(ax) + Tr(bx^{2^k+1})$ for $x \in \mathbb{F}_{2^n}^*$ has the distribution of weights and corresponding exponential sums in Table I, where $h = \frac{n-1}{2}$. For a, b, c in $\mathbb{F}_{2^n}$ for odd n, a linear cyclic subcode C_T given by $f(x) = Tr(ax) + Tr(bx^3) + Tr(cx^5)$ for $x \in \mathbb{F}_{2^n}^*$ has the distribution of weights and corresponding exponential sums in Table II. In both Tables, the exponential sum means $\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)}$ with $f(0) = 0$. There are several different ways to establish the validity of weight distribution of C_G, for example, see [3] and [11]. For those of C_T, see [11].

III. New family of binary sequences with large size

In this section, we present new families of binary sequences with large family sizes as well as large linear spans for odd and even n.

TABLE I
WEIGHT DISTRIBUTION OF C_G GIVEN BY $f(x) = Tr(ax) + Tr(bx^{2^k+1})$

Weight	Exponential Sum	Distribution
0	2^n	1
$2^{n-1} - 2^{n-h-1}$	2^{n-h}	$(2^n-1)(2^{2h-1}+2^{h-1})$
2^{n-1}	0	$(2^n-1)(2^n-2^{2h}+1)$
$2^{n-1} + 2^{n-h-1}$	-2^{n-h}	$(2^n-1)(2^{2h-1}-2^{h-1})$

TABLE II
WEIGHT DISTRIBUTION OF C_T GIVEN BY $f(x) = Tr(ax) + Tr(bx^3) + Tr(cx^5)$

Weight	Exponential Sum	Distribution
0	2^n	1
$2^{n-1} \pm 2^{\frac{n+1}{2}}$	$\mp 2^{\frac{n+3}{2}}$	$\frac{1}{3}(2^n-1)\cdot 2^{\frac{n-5}{2}}\cdot(2^{\frac{n-3}{2}}\mp 1)\cdot(2^{n-1}-1)$
2^{n-1}	0	$(2^n-1)\cdot(9\cdot 2^{2n-4}+3\cdot 2^{n-3}+1)$
$2^{n-1} \pm 2^{\frac{n-1}{2}}$	$\mp 2^{\frac{n+1}{2}}$	$\frac{1}{3}(2^n-1)\cdot 2^{\frac{n-3}{2}}\cdot(2^{\frac{n-1}{2}}\mp 1)\cdot(5\cdot 2^{n-1}+4)$

A. Construction of $\mathcal{S}_o(\rho)$ for odd n

Construction 1: For odd $n = 2l+1$ and an integer ρ with $1 \le \rho \le l$, a family $\mathcal{S}_o(\rho)$ of binary sequences is defined by

$$\mathcal{S}_o(\rho) = \{\mathbf{s}^{\mathbf{\Lambda}} \mid \mathbf{\Lambda} = (\lambda_0, \cdots, \lambda_{\rho-1}),\ \lambda_i \in \mathbb{F}_{2^n}\}$$

where $\mathbf{s}^{\mathbf{\Lambda}} = \{s_0^{\mathbf{\Lambda}}, s_1^{\mathbf{\Lambda}}, \cdots, s_{2^n-2}^{\mathbf{\Lambda}}\}$ is a binary sequence of period 2^n-1 with $s_t^{\mathbf{\Lambda}} = s_{\mathbf{\Lambda}}(\alpha^t)$ for a primitive element α of $\mathbb{F}_{2^n}$, where $s_{\mathbf{\Lambda}}(x)$, the trace representation of $s_t^{\mathbf{\Lambda}}$, is given by

$$\begin{aligned} s_{\mathbf{\Lambda}}(x) &= s_{\lambda_0,\cdots,\lambda_{\rho-1}}(x) \\ &= Tr(\lambda_0 x) + \sum_{i=1}^{\rho-1} Tr(\lambda_i x^{1+2^i}) + \sum_{i=\rho}^{l} Tr(x^{1+2^i}) \end{aligned} \quad (6)$$

for $x \in \mathbb{F}_{2^n}^*$.

The parameters of a new signal set $\mathcal{S}_o(\rho)$ are determined by the following theorem.

Theorem 1: For odd $n = 2l+1$ and an integer ρ with $1 \le \rho \le l$, the family $\mathcal{S}_o(\rho)$ has $2^{n\rho}$ cyclically distinct binary sequences of period 2^n-1. The correlation of sequences is $(2\rho+2)$-valued and maximum correlation is $1+2^{\frac{n+2\rho-1}{2}}$. Therefore, $\mathcal{S}_o(\rho)$ constitutes a $(2^n-1, 2^{n\rho}, 1+2^{\frac{n+2\rho-1}{2}})$ signal set.

In order to show Theorem 1, i.e., to determine family size and correlation of $\mathcal{S}_o(\rho)$, we need the following lemmas.

Lemma 1: All sequences in $\mathcal{S}_o(\rho)$ are cyclically distinct. Thus, the family size of $\mathcal{S}_o(\rho)$ is $2^{n\rho}$.

Proof: Consider a time shifted version of a sequence in $\mathcal{S}_o(\rho)$ represented by

$$\begin{aligned} &s_{\mathbf{\Theta}}(\delta x) \\ &= Tr(\theta_0\delta x) + \sum_{i=1}^{\rho-1} Tr(\theta_i \delta^{1+2^i} x^{1+2^i}) + \sum_{i=\rho}^{l} Tr(\delta^{1+2^l} x^{1+2^i}) \end{aligned}$$

for $\mathbf{\Theta} = (\theta_0, \cdots, \theta_{\rho-1})$, $\theta_i \in \mathbb{F}_{2^n}$, and $\delta \in \mathbb{F}_{2^n}^*$. It is identical to another sequence of (6), i.e., $s_{\mathbf{\Lambda}}(x) = s_{\mathbf{\Theta}}(\delta x)$ for all $x \in \mathbb{F}_{2^n}^*$ if and only if

$$\begin{aligned} &\lambda_0 = \theta_0\delta,\ \lambda_i = \theta_i\delta^{1+2^i} \text{ for } 1 \le i < \rho, \\ &\text{and } \delta^{1+2^i} = 1 \text{ for } \rho \le i \le l. \end{aligned} \quad (7)$$

From $\gcd(2^n-1, 1+2^{\frac{n-1}{2}}) = 1$ for odd n, $\delta = 1$ is a unique solution achieving $\delta^{1+2^l} = 1$. If $\delta = 1$ in (7), it only gives a trivial solution of $\lambda_i = \theta_i$ for $0 \le i < \rho$. Thus, sequences in $\mathcal{S}_o(\rho)$ represented by $s_{\mathbf{\Lambda}}(x)$ for any λ_i in $\mathbb{F}_{2^n}$ with $0 \le i < \rho$, are cyclically distinct. ■

The crosscorrelation of two sequences $\mathbf{s}^{\mathbf{\Lambda}}$ and $\mathbf{s}^{\mathbf{\Theta}}$ in $\mathcal{S}_o(\rho)$ is given by

$$C_{\mathbf{s}^{\mathbf{\Lambda}},\mathbf{s}^{\mathbf{\Theta}}}(\tau) = -1 + \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)}$$

where

$$f(x) = Tr(\eta_0 x) + \sum_{i=1}^{l} Tr(\eta_i x^{1+2^i}) \quad (8)$$

where

$$\eta_i = \begin{cases} \lambda_0 + \theta_0\delta, & i = 0 \\ \lambda_i + \theta_i\delta^{1+2^i}, & 1 \le i < \rho \\ 1 + \delta^{1+2^i}, & \rho \le i \le l \end{cases} \quad (9)$$

for $\lambda_i, \theta_i \in \mathbb{F}_{2^n}$ with $0 \le i < \rho$ and $\delta = \alpha^\tau \in \mathbb{F}_{2^n}^*$ where α is a primitive element of $\mathbb{F}_{2^n}$. In other words, the sum of $\mathbf{s}^{\mathbf{\Lambda}}$ and a τ-shifted version of $\mathbf{s}^{\mathbf{\Theta}}$ can be considered as a codeword given by $f(x)$ in (8). Thus, we need to investigate the exponential sum of $f(x)$ for the correlation of a pair of sequences. In the following, we classify the exponential sum in terms of values of η_i's.

Case 1. $\eta_i = 0$ for $0 \le i \le l$. In this case, $f(x) = 0$. This corresponds to a trivial exponential sum corresponding to in-phase autocorrelation of a sequence. Thus, its exponential sum is 2^n.

Case 2. $\eta_0 \ne 0$ and $\eta_i = 0$ for $1 \le i \le l$. In this case, $f(x) = Tr(\eta_0 x)$. Hence, we have the exponential sum 0 for any $\eta_0 \in \mathbb{F}_{2^n}^*$ from the orthogonality of the trace function [4].

Case 3. At least one $\eta_i \ne 0$ for $1 \le i \le l$. In this case, $f(x)$ is equivalent to a quadratic Boolean function. Thus, we need to investigate the number of roots of its symplectic form $B_f(x,z)$

in order to apply Fact 1 for determining the distribution of exponential sums of $f(x)$.

Lemma 2: For odd $n = 2l+1$ and an integer ρ with $1 \le \rho \le l$, let η_i's of $f(x)$ in (8) be given such that at least one η_i is nonzero for $1 \le i \le l$. Then, the symplectic form $B_f(x,z)$ associated with $f(x)$ has at most $2^{2\rho-1}$ roots in $x \in \mathbb{F}_{2^n}$ for all $z \in \mathbb{F}_{2^n}^*$.

Proof: For given η_i's, the symplectic form $B_f(x,z)$ associated with $f(x)$ is given by

$$\begin{aligned} B_f(x,z) &= f(x) + f(x+z) + f(z) \\ &= \sum_{i=1}^{l} Tr\left(\eta_i (x z^{2^i} + x^{2^i} z)\right) \\ &= Tr\left(z \sum_{i=1}^{l} (\eta_i^{2^{-i}} x^{2^{-i}} + \eta_i x^{2^i})\right) \\ &= Tr\left(z L(x)\right) \end{aligned} \tag{10}$$

where $L(x) = \sum_{i=1}^{l} (\eta_i^{2^{-i}} x^{2^{-i}} + \eta_i x^{2^i})$. $B_f(x,z) = 0$ for all $z \in \mathbb{F}_{2^n}^*$ if and only if $L(x) = 0$. From (9),

$$\begin{aligned} L(x) =& \sum_{i=1}^{\rho-1} (\eta_i^{2^{-i}} x^{2^{-i}} + \eta_i x^{2^i}) \\ &+ \sum_{i=\rho}^{l} ((1+\delta^{1+2^{-i}}) x^{2^{-i}} + (1+\delta^{1+2^i}) x^{2^i}) \\ =& \sum_{i=1}^{\rho-1} (\eta_i^{2^{-i}} x^{2^{-i}} + \eta_i x^{2^i}) \\ &+ \sum_{i=1}^{\rho-1} ((1+\delta^{1+2^{-i}}) x^{2^{-i}} + (1+\delta^{1+2^i}) x^{2^i}) \\ &+ \sum_{i=1}^{l} ((1+\delta^{1+2^{-i}}) x^{2^{-i}} + (1+\delta^{1+2^i}) x^{2^i}) \\ =& \sum_{i=1}^{\rho-1} ((\eta_i^{2^{-i}} + 1 + \delta^{1+2^{-i}}) x^{2^{-i}} + (\eta_i + 1 + \delta^{1+2^i}) x^{2^i}) \\ &+ \sum_{i=1}^{n-1} (1+\delta^{1+2^i}) x^{2^i}. \end{aligned}$$

Note that

$$\begin{aligned} \sum_{i=1}^{n-1} (1+\delta^{1+2^i}) x^{2^i} &= \sum_{i=1}^{n-1} x^{2^i} + \delta \sum_{i=1}^{n-1} (\delta x)^{2^i} \\ &= x + Tr(x) + \delta(Tr(\delta x) + \delta x) \\ &= (1+\delta^2) x + Tr(x) + \delta Tr(\delta x). \end{aligned} \tag{11}$$

Let

$$\gamma_i = \eta_i + 1 + \delta^{1+2^i}, \quad 1 \le i \le \rho - 1. \tag{12}$$

Then, we have $\gamma_i^{2^{-i}} = \eta_i^{2^{-i}} + 1 + \delta^{1+2^{-i}}$. Together with (11),

$$L(x) = q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x) + Tr(x) + \delta Tr(\delta x) \tag{13}$$

where

$$q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x) = (1+\delta^2) x + \sum_{i=1}^{\rho-1} (\gamma_i^{2^{-i}} x^{2^{-i}} + \gamma_i x^{2^i}). \tag{14}$$

For $L(x) = 0$, we have to count the number of solutions in the equation

$$q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x) + Tr(x) + \delta Tr(\delta x) = 0 \tag{15}$$

for given γ_i's in $\mathbb{F}_{2^n}$ and δ in $\mathbb{F}_{2^n}^*$.

Next, we verify that $q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x)$ is not a constant polynomial of x (i.e., a trivial polynomial). If $\delta = 1$, $\gamma_i = \eta_i$ for $1 \le i \le \rho - 1$ from (12), and $\eta_i = 0$ for $\rho \le i \le l$ from (9). Then, at least one γ_i is nonzero for $1 \le i \le \rho - 1$ because at least one η_i is nonzero. If $\delta \ne 1$, on the other hand, $(1+\delta^2)$ cannot be zero although γ_i may be zero for all $1 \le i \le \rho - 1$. Therefore, $q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x)$ is a polynomial of x with at least one nonzero coefficient of x^{2^i} for $-\rho < i < \rho$.

For the nontrivial polynomial $q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x)$, the equation in (15) can be divided into four classes.

i) $Tr(x) = 0$ and $Tr(\delta x) = 0 \Rightarrow q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x) = 0$,
ii) $Tr(x) = 0$ and $Tr(\delta x) = 1 \Rightarrow q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x) + \delta = 0$,
iii) $Tr(x) = 1$ and $Tr(\delta x) = 0 \Rightarrow q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x) + 1 = 0$,
iv) $Tr(x) = 1$ and $Tr(\delta x) = 1 \Rightarrow q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x) + 1 + \delta = 0$.

Thus, the left-hand side of (15) can be presented by

$$\begin{aligned} A_a(x) &= q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x) + a \\ &= a + (1+\delta^2) x + \sum_{i=1}^{\rho-1} (\gamma_i^{2^{-i}} x^{2^{-i}} + \gamma_i x^{2^i}) \end{aligned}$$

for $a \in \{0, 1, \delta, 1+\delta\}$. Then,

$$\begin{aligned} A_a^{2^{\rho-1}}(x) &= \left[q_{\gamma_1, \cdots, \gamma_{\rho-1}, \delta}(x) + a\right]^{2^{\rho-1}} \\ &= a^{2^{\rho-1}} + (1+\delta^2)^{2^{\rho-1}} x^{2^{\rho-1}} \\ &\quad + \sum_{i=1}^{\rho-1} (\gamma_i^{2^{\rho-1-i}} x^{2^{\rho-1-i}} + \gamma_i^{2^{\rho-1}} x^{2^{\rho-1+i}}) \end{aligned}$$

and

$$A_a^{2^{\rho-1}}(x) = 0 \iff A_a(x) = 0.$$

Thus, solutions for $A_a^{2^{\rho-1}}(x) = 0$ are all solutions for $A_a(x) = 0$, and vice versa. Therefore, the number of solutions for $A_a(x) = 0$ is equal to that for $A_a^{2^{\rho-1}}(x) = 0$. Since the maximum degree of $A_a^{2^{\rho-1}}(x)$ is $2^{2(\rho-1)}$, $A_a^{2^{\rho-1}}(x) = 0$ has at most $2^{2(\rho-1)}$ solutions. If $\delta \ne 1$, the solutions of $A_a^{2^{\rho-1}}(x) = 0$ are disjoint for different $a \in \{0, 1, \delta, 1+\delta\}$, so the total number of solutions of (15) is at most $2^{2(\rho-1)} \cdot 4 = 2^{2\rho}$. If $\delta = 1$, on the other hand, the solutions of $A_a^{2^{\rho-1}}(x) = 0$ are disjoint for different $a \in \{0, 1\}$, so the total number of solutions of (15) is at most $2^{2(\rho-1)} \cdot 2 = 2^{2\rho-1}$. From Fact 1, meanwhile, a possible number of roots of $B_f(x,z)$ is 2^{n-2h} where $n-2h$ is a positive odd integer for odd n. For any value of δ in $\mathbb{F}_{2^n}^*$, therefore, the maximum number of solutions of $B_f(x,z) = 0$ is $2^{2\rho-1}$. ■

From Lemma 2, we see that $B_f(x,z) = 0$ has 2^{n-2h} solutions for an integer h where $n-2h$ is a positive odd integer less than or equal to $2\rho-1$. From Fact 1, for all $\eta_0 \in \mathbb{F}_{2^n}$, the exponential sum of $f(x)$ can take on values of 0 and $\pm 2^{n-h}$ for an integer h where $n-2h$ is a positive odd integer less than or equal to $2\rho-1$. At this point, we need to show that it can take on values of 0 and $\pm 2^{n-h}$ for all h's such that $n-2h$

is *every* positive odd integer less than or equal to $2\rho-1$. In order to do so, we need the following fact from [11]. (Note. we slightly change the representation of the result in [11].)

Fact 2 ([11], page 454): Let ζ be a set of symplectic forms of (10). For some fixed integer d with $1 \le d \le \lfloor \frac{n}{2} \rfloor$, assume that the rank of every nonzero form in ζ is at least $2d$. Then, the maximum size of ζ is given by

$$|\zeta|_{\max} = \begin{cases} 2^{n(\frac{n+1}{2}-d)} & \text{for odd } n \\ 2^{(n-1)(\frac{n+2}{2}-d)} & \text{for even } n. \end{cases}$$

In the following, we consider a specific linearized polynomial and investigate a rank of the symplectic form corresponding to the polynomial.

Lemma 3: For an integer $\rho > 1$, consider a symplectic form $B(x,z) = Tr(zU(x))$ where a linearized polynomial $U(x)$ is given by

$$U(x) = \sum_{i=1}^{\rho-1} (\gamma_i^{2^{-i}} x^{2^{-i}} + \gamma_i x^{2^i})$$

where γ_i can be any element in $\mathbb{F}_{2^n}$. For odd n, the rank of $B(x,z)$ is at least $n-2\rho+3$ and every possible rank $2h$ for $n-2\rho+3 \le 2h \le n-1$ occurs at least once when γ_i runs through $\mathbb{F}_{2^n}$. For even n, on the other hand, the rank is at least $n-2\rho+2$ and every possible rank $2h$ for $n-2\rho+2 \le 2h \le n-2$ occurs at least once when γ_i runs through $\mathbb{F}_{2^n}$.

Proof: Note that the odd case in Lemma 3 is implicitly known from Theorem 16 and Corollary 17 of Chapter 15 in [11]. Here, we reproduce it for completeness.

Let k be an integer with $2 \le k \le \rho$. By $\gamma_k = \gamma_{k+1} = \cdots = \gamma_{\rho-1} = 0$, we have $B_k(x,z) = Tr(zU_k(x))$ where $U_k(x) = \sum_{i=1}^{k-1} (\gamma_i^{2^{-i}} x^{2^{-i}} + \gamma_i x^{2^i})$. Then, the maximum degree of $(U_k(x))^{2^{(k-1)}}$ is $2^{2(k-1)}$ and thus the rank of $B_k(x,z)$ is at least $n-2k+2$ (If n is odd, the rank is at least $n-2k+3$ because it should be even).

For odd n, assume that the rank $n-2k+3$ never occurs for all γ_i's with $1 \le i \le k-1$, and thus the rank of $B_k(x,z)$ is at least $n-2k+5$. Then, from Fact 2, the maximum size of a set ζ of such $B_k(x,z)$'s is $|\zeta|_{\max} = 2^{n(k-2)}$. However, its actual size is $2^{n(k-1)}$ when γ_i's with $1 \le i \le k-1$ run through $\mathbb{F}_{2^n}$, which is greater than $|\zeta|_{\max}$. Therefore, the rank of $n-2k+3$ occurs at least once when γ_i's run through $\mathbb{F}_{2^n}$. For even n, if the rank $n-2k+2$ never occurs, then the minimum rank is $n-2k+4$. Similarly, the maximum size of a set ζ of such $B_k(x,z)$'s is $|\zeta|_{\max} = 2^{(n-1)(k-1)}$, which is smaller than its actual size $2^{n(k-1)}$. Therefore, the rank of $n-2k+2$ occurs at least once when γ_i's run through $\mathbb{F}_{2^n}$.

Since $U(x)$ contains every $U_k(x)$'s for $2 \le k \le \rho$ when γ_i's with $1 \le i \le \rho-1$ run through $\mathbb{F}_{2^n}$, the assertion of Lemma 3 follows. ∎

Using Lemma 3, we have the following result.

Lemma 4: For odd $n = 2l+1$ and an integer ρ with $1 \le \rho \le l$, let η_i's of $f(x)$ in (8) be given such that at least one η_i is nonzero for $1 \le i \le l$, and $\eta_0 \in \mathbb{F}_{2^n}$. Then, the exponential sum of $f(x)$ can take on values of 0 and $\pm 2^{\frac{n-h}{2}}$ for an integer h where $n-2h$ is every positive odd integer less than or equal to $2\rho-1$.

Proof: For $f(x)$, by $\delta = 1$, then the linearized polynomial given in (14) (if $\delta = 1$, it is identical to (13)) is of the form of $U(x)$ in Lemma 3. From Lemma 3, the exponential sum of $f(x)$ for $\delta = 1$ is equal to 0 or $\pm 2^{\frac{n-h}{2}}$ for all h's such that $n-2h = 1, 3, \cdots, 2\rho-3$. Thus, we can say that the exponential sum of $f(x)$ takes on values of 0 and $\pm 2^{\frac{n-h}{2}}$ for every integer h such that $n-2h = 1, 3, \cdots, 2\rho-3$.

Now, it suffices to show that the exponential sum takes on values of $\pm 2^{\frac{n-h}{2}}$ at least once when $n-2h = 2\rho-1$. Or equivalently, we need to show the symplectic form $B_f(x,z)$ in (10) has a rank of $n-2\rho+1$ at least once for some η_i's. Assume that this rank never occurs for all η_i's. Then, the rank of $B_f(x,z)$ is at least $2d = n-2\rho+3$. From Fact 2, the maximum size of a set ζ of such $B_f(x,z)$'s is $|\zeta|_{\max} = 2^{n(\rho-1)}$. However, from (9) and (10), its actual size is $2^{n\rho}$, which is greater than $|\zeta|_{\max}$. Therefore, the rank of $n-2\rho+1$ occurs at least once when η_i's run through $\mathbb{F}_{2^n}$. This completes the proof of Lemma 4. ∎

Combining the cases 1, 2, and 3, we have the following result on the correlation of sequences in $\mathcal{S}_o(\rho)$.

Lemma 5: The correlation of binary sequences in $\mathcal{S}_o(\rho)$ is $(2\rho+2)$-valued and maximum correlation $C_{\max}$ is $1+2^{\frac{n+2\rho-1}{2}}$.

Proof: For a trace representation $s_{\mathbf{\Lambda}}(x)$ of each sequence $\mathbf{s}^{\mathbf{\Lambda}}$ in $\mathcal{S}_o(\rho)$, we can consider the exponential sum of $f(x)$ in (8). In cases 1 and 2, the exponential sum has 2^n and 0 values. For all η_i's in case 3, from Lemma 4, the exponential sum takes on 0 and $\pm 2^{\frac{n-h}{2}}$ for each integer h such that $n-2h = 1, 3, \cdots, 2\rho-1$. Thus, the exponential sum takes on 2ρ nonzero distinct values. Including 2^n and 0, therefore, the overall exponential sum is $(2\rho+2)$-valued. Equivalently, the correlation of sequences in $\mathcal{S}_o(\rho)$ is $(2\rho+2)$-valued. Since maximum value for $n-2h$ is determined by $2\rho-1$ from Lemma 4, $C_{\max} = |-1-2^{\frac{n-h}{2}}| = 1+2^{\frac{n+2\rho-1}{2}}$. ∎

Proof of Theorem 1. The results follow directly from Lemmas 1 and 5. ∎

Remark 1: If $\rho = 1$,

$$s_{\lambda_0}(x) = Tr(\lambda_0 x) + \sum_{i=1}^{l} Tr(x^{1+2^i}), \quad x \in \mathbb{F}_{2^n}^* \tag{16}$$

which represents the Gold-like Sequences introduced by Boztas and Kumar [1]. From Lemma 4, a positive odd integer less than 2ρ is only 1, so $n-2h = 1$. Hence, $h = \frac{n-1}{2}$. Finally, Gold-like sequences given by (16) have four-valued correlation, or $\{2^n-1, -1, -1 \pm 2^{\frac{n+1}{2}}\}$ for all $\lambda_0 \in \mathbb{F}_{2^n}$.

Example 1: If $\rho = 2$,

$$s_{\lambda_0,\lambda_1}(x) = Tr(\lambda_0 x) + Tr(\lambda_1 x^3) + \sum_{i=2}^{l} Tr(x^{1+2^i}) \tag{17}$$

for $x \in \mathbb{F}_{2^n}^*$ and odd $n \ge 5$. From Lemma 4, $n-2h$ is a positive odd integer less than 4, so $n-2h = 1$ and 3. Thus, $h = \frac{n-1}{2}$ and $\frac{n-3}{2}$. From Lemma 5, sequences given by (17) have

six-valued correlation, or $\{2^n-1, -1, -1\pm 2^{\frac{n+1}{2}}, -1\pm 2^{\frac{n+3}{2}}\}$ for all $\lambda_0, \lambda_1 \in \mathbb{F}_{2^n}$. For sequences represented by $s_{\lambda_0,\lambda_1}(x)$ and $s_{\theta_0,\theta_1}(\delta x)$, the corresponding $f(x)$ in (8) constitutes a linear cyclic subcode with five nonzero distinct weights if we assume that δ can be any element in $\mathbb{F}_{2^n}$. In fact, the correlation distribution of $\mathcal{S}_o(2)$ can be derived from Table II. The details will be discussed in Section IV.

Example 2: If $\rho = 3$,

$$s_{\lambda_0,\lambda_1,\lambda_2}(x) = Tr(\lambda_0 x) + Tr(\lambda_1 x^3) + Tr(\lambda_2 x^5) + \sum_{i=3}^{l} Tr(x^{1+2^i}) \quad (18)$$

for $x \in \mathbb{F}_{2^n}^*$ and odd $n \geq 7$. Similarly, $h = \frac{n-1}{2}$, $\frac{n-3}{2}$, and $\frac{n-5}{2}$. From Lemma 5, sequences given by (18) have eight-valued correlation, or $\{2^n-1, -1, -1\pm 2^{\frac{n+1}{2}}, -1\pm 2^{\frac{n+3}{2}}, -1\pm 2^{\frac{n+5}{2}}\}$ for all $\lambda_0, \lambda_1, \lambda_2 \in \mathbb{F}_{2^n}$.

B. Construction of $\mathcal{S}_e(\rho)$ for even n

Construction 2: For even $n = 2l$ and an integer ρ with $1 \leq \rho < l$, a family $\mathcal{S}_e(\rho)$ of binary sequences is defined by

$$\mathcal{S}_e(\rho) = \{\mathbf{s}^{\mathbf{\Lambda}} \mid \mathbf{\Lambda} = (\lambda_0, \cdots, \lambda_{\rho-1}),\ \lambda_i \in \mathbb{F}_{2^n}\}$$

where $\mathbf{s}^{\mathbf{\Lambda}} = \{s_0^{\mathbf{\Lambda}}, s_1^{\mathbf{\Lambda}}, \cdots, s_{2^n-2}^{\mathbf{\Lambda}}\}$ is a binary sequence of period 2^n-1 with $s_t^{\mathbf{\Lambda}} = s_{\mathbf{\Lambda}}(\alpha^t)$ for a primitive element α of $\mathbb{F}_{2^n}$, where $s_{\mathbf{\Lambda}}(x)$, the trace representation of $s_t^{\mathbf{\Lambda}}$, is given by

$$s_{\mathbf{\Lambda}}(x) = Tr(\lambda_0 x) + \sum_{i=1}^{\rho-1} Tr(\lambda_i x^{1+2^i}) + \sum_{i=\rho}^{l-1} Tr(x^{1+2^i}) + Tr_1^l(x^{1+2^l}) \quad (19)$$

for $x \in \mathbb{F}_{2^n}^*$.

Theorem 2: For even $n = 2l$ and an integer ρ with $1 \leq \rho < l$, the family $\mathcal{S}_e(\rho)$ has $2^{n\rho}$ cyclically distinct binary sequences of period 2^n-1. The correlation of sequences is $(2\rho+4)$-valued and maximum correlation is $1+2^{\frac{n}{2}+\rho}$. Therefore, $\mathcal{S}_e(\rho)$ constitutes a $(2^n-1, 2^{n\rho}, 1+2^{\frac{n}{2}+\rho})$ signal set.

In order to prove Theorem 2, we need the following lemmas. Since their proofs are similar to those for $\mathcal{S}_o(\rho)$, we omit the details.

Lemma 6: All sequences in $\mathcal{S}_e(\rho)$ are cyclically distinct. Hence, the family size of $\mathcal{S}_e(\rho)$ is $2^{n\rho}$.

Proof: Similar to the proof of Lemma 1, $s_{\mathbf{\Lambda}}(x) = s_{\mathbf{\Theta}}(\delta x)$ for all x in $\mathbb{F}_{2^n}^*$ if and only if (7) is achieved. If $\gcd(1+2^i, 2^n-1) = d > 1$, then d is not a factor of $\gcd(1+2^{i-1}, 2^n-1)$ since $\gcd(1+2^i, 1+2^{i-1}) = 1$ for any integer i. Hence, if $\delta_1 \neq 1$ is a solution of $\delta^{1+2^i} = 1$, then it cannot be a solution of $\delta^{1+2^{i-1}} = 1$. Meanwhile, we have at least two equations of $\delta^{1+2^i} = 1$, for $\rho \leq i \leq l$ in (7) because $\rho < l$ for even n. Thus, (7) has a unique solution given by $\delta = 1$. Hence, all sequences in $\mathcal{S}_e(\rho)$ are cyclically distinct. ∎

To investigate the correlation of sequences in $\mathcal{S}_e(\rho)$, we need to consider $f(x)$ given by

$$f(x) = Tr(\eta_0 x) + \sum_{i=1}^{l-1} Tr(\eta_i x^{1+2^i}) + Tr_1^l(\eta_l x^{1+2^l}) \quad (20)$$

where

$$\eta_i = \begin{cases} \lambda_0 + \theta_0\delta, & i = 0 \\ \lambda_i + \theta_i\delta^{1+2^i}, & 1 \leq i < \rho \\ 1 + \delta^{1+2^i}, & \rho \leq i \leq l \end{cases} \quad (21)$$

for $\lambda_i, \theta_i \in \mathbb{F}_{2^n}$ with $0 \leq i < \rho$ and $\delta \in \mathbb{F}_{2^n}^*$. If $\eta_i = 0$ for all $1 \leq i \leq l$, the exponential sum of $f(x)$ has a value of 0 or 2^n. Otherwise, we have to consider the number of solutions of $B_f(x,z) = 0$ in order to derive the distribution of exponential sums of $f(x)$ in (20).

Lemma 7: For even $n = 2l$ and an integer ρ with $1 \leq \rho < l$, let η_i's of $f(x)$ in (20) be given such that at least one η_i is nonzero for $1 \leq i \leq l$. Then, the symplectic form $B_f(x,z)$ associated with $f(x)$ has at most $2^{2\rho}$ roots in $x \in \mathbb{F}_{2^n}$ for all $z \in \mathbb{F}_{2^n}^*$.

Proof: We have

$$B_f(x,z) = Tr\left(z\left(\sum_{i=1}^{l-1}(\eta_i^{2^{-i}} x^{2^{-i}} + \eta_i x^{2^i}) + \eta_l x^{2^l}\right)\right) = Tr\left(zL(x)\right).$$

Thus, $B_f(x,z) = 0$ for all $z \in \mathbb{F}_{2^n}^*$ if and only if $L(x) = 0$. Using the same approach as we did in the proof of Lemma 2, we obtain $L(x)$ identical to (13) and the equation (15). Following the same steps as in the proof of Lemma 2, we can derive that the number of solutions of $B_f(x,z) = 0$ is at most $2^{2\rho}$. ∎

Lemma 8: For even $n = 2l$ and an integer ρ with $1 \leq \rho < l$, let η_i's of $f(x)$ in (20) be given such that at least one η_i is nonzero for $1 \leq i \leq l$, and $\eta_0 \in \mathbb{F}_{2^n}$. Then, the exponential sum of $f(x)$ can take on values of 0 and $\pm 2^{n-h}$ for an integer h where $n-2h$ is zero or every positive even integer less than or equal to 2ρ. Hence, the correlation of binary sequences in $\mathcal{S}_e(\rho)$ is $(2\rho+4)$-valued and maximum correlation $C_{\max}$ is $1+2^{\frac{n}{2}+\rho}$.

Proof: For even n, we see that while each η_i with $1 \leq i < \rho$ runs through $\mathbb{F}_{2^n}$, each γ_i of $L(x)$ in (13) runs through $\mathbb{F}_{2^n}$. From Lemma 3, therefore, we obtain that the exponential sum of $f(x)$ takes on values of 0 and $\pm 2^{n-h}$ for every integer h such that $n-2h = 2, 4, \cdots, 2\rho-2$.

Next, we will show that the cases $n-2h = 0$ and $n-2h = 2\rho$ also occur. If all γ_i's of $L(x)$ in (13) are zero, then $L(x) = 0$ has at most four solutions, i.e., $x = 0, \frac{\delta}{1+\delta^2}, \frac{1}{1+\delta^2}$, and $\frac{1}{1+\delta}$ where the last three solutions are valid only if $Tr\left(\frac{1}{1+\delta}\right) = 1$. Thus, we have a unique solution $x = 0$ for δ such that $Tr\left(\frac{1}{1+\delta}\right) = 0$. Hence, the case $n-2h = 0$ occurs at least once. Applying Fact 2, we also obtain that the exponential sum of $f(x)$ can take on values of $\pm 2^{n-h}$ at least once such that $n-2h = 2\rho$ in the similar way to the proof of Lemma 4.

Therefore, the exponential sum of $f(x)$ takes on values of 0 and $\pm 2^{n-h}$ for an integer h where $n-2h = 0, 2, \cdots, 2\rho$. Including 0 and 2^n, the correlation is $(2\rho+4)$-valued. Furthermore, $C_{\max} = 1 + 2^{\frac{n}{2}+\rho}$ for $h = \frac{n}{2} - \rho$. ■

Proof of Theorem 2. The results follow directly from Lemma 6 and Lemma 8. ■

Remark 2: If $\rho = 1$,

$$s_{\lambda_0}(x) = Tr(\lambda_0 x) + \sum_{i=1}^{l-1} Tr(x^{1+2^i}) + Tr_1^l(x^{1+2^l}), \quad x \in \mathbb{F}_{2^n}^* \tag{22}$$

which represents the sequences constructed by Udaya [19]. From Lemma 8, $n-2h = 0$ and 2. Hence, $h = \frac{n}{2}$ and $\frac{n}{2}-1$. Finally, the sequences given by (22) has six-valued correlation, or $\{2^n-1, -1, -1 \pm 2^{\frac{n}{2}}, -1 \pm 2^{\frac{n}{2}+1}\}$ for all $\lambda_0 \in \mathbb{F}_{2^n}$.

Example 3: If $\rho = 2$,

$$s_{\lambda_0,\lambda_1}(x) = Tr(\lambda_0 x) + Tr(\lambda_1 x^3) + \sum_{i=2}^{l} Tr(x^{1+2^i}) + Tr_1^l(x^{1+2^l}) \tag{23}$$

for $x \in \mathbb{F}_{2^n}^*$ and even $n \geq 6$. From Lemma 8, $n-2h = 0, 2, 4$ and thus, $h = \frac{n}{2}, \frac{n}{2}-1, \frac{n}{2}-2$. Hence, the correlation of sequences given by (23) belongs to $\{2^n-1, -1, -1 \pm 2^{\frac{n}{2}}, -1 \pm 2^{\frac{n}{2}+1}, -1 \pm 2^{\frac{n}{2}+2}\}$ for all $\lambda_0, \lambda_1 \in \mathbb{F}_{2^n}$.

Remark 3: Contrary to odd n, $f(x)$ in (20) does not constitute a linear cyclic subcode because η_i with $\rho \leq i < l$ may not run through all elements in $\mathbb{F}_{2^n}$. Therefore, we should use different methods from odd n to investigate the correlation distribution in $\mathcal{S}_e(2)$, which is a problem we are working on.

C. Linear spans of $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$

The linear spans of binary sequences in $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ are determined by the number of nonzero λ_i's with $0 \leq i < \rho$.

Theorem 3: In the family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$), consider a sequence represented by $s_{\mathbf{\Lambda}}(x)$ where j λ_i's in $\mathbf{\Lambda} = (\lambda_0, \cdots, \lambda_{\rho-1})$ are equal to 0. Let $LS_j(\rho)$ be the linear span of the sequence. Then,

$$LS_j(\rho) = \frac{n(n-2j+1)}{2}, \quad 0 \leq j \leq \rho$$

and there are $\binom{\rho}{j} \cdot (2^n-1)^{\rho-j}$ sequences in $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) having linear span $LS_j(\rho)$. From this result, the linear span of sequences in $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) and its distribution are shown in Table III.

Proof: Firstly, consider the linear span of sequences in $\mathcal{S}_o(\rho)$. In Construction 1, a sequence represented by $s_{\mathbf{\Lambda}}(x)$ has a total $l+1 = \frac{n+1}{2}$ trace terms and each trace term has the linear span of n. If j λ_i's of the sequence are equal to 0, it has $(\frac{n+1}{2} - j)$ nonzero trace terms and the corresponding linear span of the sequence is given by

$$LS_j(\rho) = \left(\frac{n+1}{2} - j\right) \cdot n = \frac{n(n-2j+1)}{2}, \quad 0 \leq j \leq \rho.$$

TABLE III

LINEAR SPAN AND THE NUMBER OF CORRESPONDING SEQUENCES IN $\mathcal{S}_o(\rho)$ (OR $\mathcal{S}_e(\rho)$)

$LS(\rho)$	Number of sequences
$\frac{n(n+1)}{2}$	$(2^n-1)^\rho$
$\frac{n(n-1)}{2}$	$\binom{\rho}{1} \cdot (2^n-1)^{\rho-1}$
$\frac{n(n-3)}{2}$	$\binom{\rho}{2} \cdot (2^n-1)^{\rho-2}$
$\vdots$	$\vdots$
$\frac{n(n-2\rho+1)}{2}$	1

Since j λ_i's are 0 and $(\rho-j)$ λ_i's are nonzero, the number of corresponding sequences given by $s_{\mathbf{\Lambda}}(x)$ is $\binom{\rho}{j} \cdot (2^n-1)^{\rho-j}$. Applying this result to each j with $0 \leq j \leq \rho$, we obtain the linear span $LS(\rho)$ of binary sequences in $\mathcal{S}_o(\rho)$ with the distribution shown in Table III. Using the similar approach to odd case, we see that the linear span of sequences in $\mathcal{S}_e(\rho)$ is the same as $\mathcal{S}_o(\rho)$. ■

Corollary 1: The maximum and minimum linear spans of sequences in $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ are $\frac{n(n+1)}{2}$ and $\frac{n(n-2\rho+1)}{2}$, respectively.

D. Comparison of families of binary sequences

In Table IV, we give some comparisons of the families $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ with well known families of binary sequences with low correlation. In Table IV, we include a family of binary sequences constructed from Chang *et al.*'s investigation of a binary cyclic code based on three-term sequences [2]. For odd $n = 2l+1$, each sequence in the family is represented by

$$s_{\lambda_0,\lambda_1}(x) = Tr(\lambda_0 x) + Tr(\lambda_1 x^r) + Tr(x^{r^2}), \quad r = 2^{l+1}+1$$

for $x \in \mathbb{F}_{2^n}^*$ and all λ_0, λ_1 in $\mathbb{F}_{2^n}$. On the other hand, a family of sequences defined by $s_{\lambda_0,\lambda_1}(x) = Tr(\lambda_0 x) + Tr(\lambda_1 x^3) + Tr(x^5)$ from the dual of the triple error correcting BCH code, has the same period, family size, $C_{\max}$, and maximum linear span as those of Chang *et al.*'s for all λ_0, λ_1 in $\mathbb{F}_{2^n}$ for odd n [11].

In terms of periods and maximum linear spans, new families $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ are identical to the families of Gold-like sequences and Udaya's, respectively. At the expense of maximum correlation $C_{\max}$, however, we have larger family sizes in $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ than in those families. Basically, we obtain $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ with exponentially increased family sizes by the linear increase of $C_{\max}$ from optimal correlation. Furthermore, we can choose ρ and the corresponding family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) for its specific application. For example, if low correlation is more crucial than a large family size in the application, then a small value of ρ is chosen. Otherwise, we choose a large value of ρ in order to get a large family size. Considering the flexibility due to ρ, $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ have larger family sizes than any other families of binary sequences in Table IV.

In $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$, m-sequences are not included for increasing their minimum linear spans. Thus, they have larger minimum linear spans than any other binary sequence families in Table IV except for bent sequences, which will be good property for potential cryptographic applications.

TABLE IV

COMPARISON OF THE FAMILIES OF BINARY SEQUENCES WITH LOW CORRELATION

Family of Sequences	Period	Family Size	$C_{\max}$	Linear Span (Maximum, Minimum)	n
Gold [3]	2^n-1	2^n+1	$1+2^{\frac{n+1}{2}}$	$(2n,n)$	odd
Kasami (Small Set) [6]	2^n-1	$2^{\frac{n}{2}}$	$1+2^{\frac{n}{2}}$	$(\frac{3n}{2},n)$	even
Kasami (Large Set) [6]	2^n-1	$2^{\frac{n}{2}}(2^n+1)-1$ or $2^{\frac{n}{2}}(2^n+1)$	$1+2^{\frac{n}{2}+1}$	$(\frac{5n}{2},n)$	even
Bent [14]	2^n-1	$2^{\frac{n}{2}}$	$1+2^{\frac{n}{2}}$	$(l_{\max},l_{\min})$ [1]	even
Boztas and Kumar [1]	2^n-1	2^n+1	$1+2^{\frac{n+1}{2}}$	$\left(\frac{n(n+1)}{2},n\right)$	odd
Udaya [19]	2^n-1	2^n+1	$1+2^{\frac{n}{2}+1}$	$\left(\frac{n(n+1)}{2},n\right)$	even
Chang *et al.* (triplet) [2] [11]	2^n-1	2^{2n}	$1+2^{\frac{n+3}{2}}$	$(3n,n)$	odd
Rothaus [15]	2^n-1	$2^{2n}+2^n+1$	$1+2^{\frac{n+3}{2}}$	$(3n,n)$	odd
Kerdock [16]	$2(2^n-1)$	2^{n-1}	$2+2^{\frac{n+1}{2}}$	$-$ [2]	odd
Delsarte-Goethals [16]	$2(2^n-1)$	2^{2n-1}	$2+2^{\frac{n+3}{2}}$	$-$ [2]	odd
New Family $\mathcal{S}_o(2)$	2^n-1	2^{2n}	$1+2^{\frac{n+3}{2}}$	$\left(\frac{n(n+1)}{2},\frac{n(n-3)}{2}\right)$	odd
New Family $\mathcal{S}_e(2)$	2^n-1	2^{2n}	$1+2^{\frac{n}{2}+2}$	$\left(\frac{n(n+1)}{2},\frac{n(n-3)}{2}\right)$	even
New Family $\mathcal{S}_o(\rho), 1<\rho\le\frac{n-1}{2}$	2^n-1	$2^{n\rho}$	$1+2^{\frac{n+2\rho-1}{2}}$	$\left(\frac{n(n+1)}{2},\frac{n(n-2\rho+1)}{2}\right)$	odd
New Family $\mathcal{S}_e(\rho), 1<\rho<\frac{n}{2}$	2^n-1	$2^{n\rho}$	$1+2^{\frac{n}{2}+\rho}$	$\left(\frac{n(n+1)}{2},\frac{n(n-2\rho+1)}{2}\right)$	even

1) $l_{\max}=\sum_{i=1}^{l-1}\binom{n}{i}+\binom{m}{l}2^l-\sum_{i=1}^{\lfloor\frac{l-1}{2}\rfloor}\binom{m}{i}$, $l_{\min}=\binom{m}{l}2^l+\sum_{i=2}^{l-1}\binom{m}{i}2^{i-1}+n$, $n=2m=4l$. See [9].
2) The linear spans of Kerdock and Delsarte-Goethals sequences are not given in [16].

IV. DISTRIBUTION OF CORRELATION VALUES OF $\mathcal{S}_o(2)$

In many applications, a family of binary sequences with both low correlation and large family size is required. Thus, we need to consider a family of sequences which can be a good compromise between correlation and family size. As a good candidate of such a family, we extensively discuss a new sequence family $\mathcal{S}_o(2)$ by analyzing its correlation distribution.

For odd $n=2l+1$, a family $\mathcal{S}_o(2)$ of binary sequences is given by

$$\mathcal{S}_o(2)=\{\mathbf{s}^{\mathbf{\Lambda}} \mid \mathbf{\Lambda}=(\lambda_0,\lambda_1),\ \lambda_i\in\mathbb{F}_{2^n}\}$$

where $\mathbf{s}^{\mathbf{\Lambda}}=\{s_0^{\mathbf{\Lambda}},s_1^{\mathbf{\Lambda}},\cdots,s_{2^n-2}^{\mathbf{\Lambda}}\}$ is a binary sequence of period 2^n-1 with $s_t^{\mathbf{\Lambda}}=s_{\lambda_0,\lambda_1}(\alpha^t)$ for a primitive element α of $\mathbb{F}_{2^n}$, where $s_{\lambda_0,\lambda_1}(x)$ is given by

$$s_{\lambda_0,\lambda_1}(x)=Tr(\lambda_0 x)+Tr(\lambda_1 x^3)+\sum_{i=2}^{l}Tr(x^{1+2^i})$$

for $x\in\mathbb{F}_{2^n}^*$.

From Theorem 1, we know that the family $\mathcal{S}_o(2)$ has 2^{2n} cyclically distinct binary sequences of period 2^n-1. The correlation of sequences in $\mathcal{S}_o(2)$ is six-valued and its maximum is $1+2^{\frac{n+3}{2}}$. Consequently, $\mathcal{S}_o(2)$ constitutes a $(2^n-1,2^{2n},1+2^{\frac{n+3}{2}})$ signal set. This data of $\mathcal{S}_o(2)$ is listed in Table IV.

Next, we will investigate the distribution of correlation values of sequences in $\mathcal{S}_o(2)$. The correlation of a pair of sequences in $\mathcal{S}_o(2)$ is derived from the exponential sum of

$$f(x)=Tr((\lambda_0+\theta_0\delta)x)+Tr((\lambda_1+\theta_1\delta^3)x^3)+\sum_{i=2}^{l}Tr((1+\delta^{1+2^i})x^{1+2^i})$$

for $\lambda_0,\lambda_1,\theta_0,\theta_1\in\mathbb{F}_{2^n}$ and $\delta\in\mathbb{F}_{2^n}^*$. With $a=\lambda_0+\theta_0\delta$, $b=\lambda_1+\theta_1\delta^3$, and $c_i=1+\delta^{1+2^i}$ for $2\le i\le l$, $f(x)$ has a form of

$$f(x)=Tr(ax)+Tr(bx^3)+\sum_{i=2}^{l}Tr(c_i x^{1+2^i}) \qquad (24)$$

for $a,b\in\mathbb{F}_{2^n}$ and $c_i\in\mathbb{F}_{2^n}\setminus\{1\}$. Depending on c_i's, the exponential sum of $f(x)$ can be classified into two exclusive cases. To facilitate the analysis, assume that c_i can be any element in $\mathbb{F}_{2^n}$.

Case 1. $c_i=0$ for $2\le i\le l$. In this case, $f(x)=Tr(ax)+Tr(bx^3)$ constitutes a linear cyclic subcode of $R(2,n)^*$ for $a,b\in\mathbb{F}_{2^n}$. Thus, the distribution of the exponential sums of $f(x)$ is identical to Table I with $h=h_1=\frac{n-1}{2}$.

Case 2. At least one $c_i\neq 0$ for $2\le i\le l$. In this case, the distribution of the exponential sums of $f(x)$ follows from the following lemma.

Lemma 9: For given c_i's with at least one nonzero c_i, $f(x)$ in (24) has five-valued exponential sum for all a,b in $\mathbb{F}_{2^n}$, and the distribution is

$$\sum_{x\in\mathbb{F}_{2^n}}(-1)^{f(x)}=\begin{cases}\pm 2^{n-h_1}, & d_1(2^{2h_1-1}\pm 2^{h_1-1}) \text{ times}\\ 0, & d_1(2^n-2^{2h_1})+d_2(2^n-2^{2h_2}) \text{ times}\\ \pm 2^{n-h_2}, & d_2(2^{2h_2-1}\pm 2^{h_2-1}) \text{ times}\end{cases} \qquad (25)$$

where $h_1=\frac{n-1}{2}$, $h_2=\frac{n-3}{2}$, and d_1,d_2 are integers such that $d_1+d_2=2^n$.

Proof: If b is fixed to an element in $\mathbb{F}_{2^n}$, the exponential sum of $f(x)$ follows the distribution in Fact 1 depending on

the rank of its symplectic form. Since $\rho = 2$, it is obvious that the symplectic form of $f(x)$ with at least one nonzero c_i may have a pair of distinct ranks $2h_1$ and $2h_2$ depending on b, where $n - 2h_1 = 1$ and $n - 2h_2 = 3$, respectively. For a subset Ω of $\mathbb{F}_{2^n}$, assume that if $b \in \Omega$, $f(x)$ has the rank $2h_1$ and otherwise, $f(x)$ has the rank $2h_2$, where $|\Omega| = d_1$ and $|\Omega^c| = d_2 = 2^n - d_1$. If b runs through all elements in $\mathbb{F}_{2^n}$, the overall distribution of the exponential sums becomes (25) after summing up the distributions for h_1 and h_2. ■

By combining both cases 1 and 2, we see that $f(x)$ has six-valued exponential sum for any c_i in $\mathbb{F}_{2^n}$. Indeed, $f(x)$ equivalently constitutes a linear cyclic subcode of $R(2,n)^*$ with five nonzero distinct weights for a, b, and c_i's in $\mathbb{F}_{2^n}$ with the assumption that each c_i determined by δ can be any element in $\mathbb{F}_{2^n}$. Hence, each codeword in the subcode has length $2^n - 1$ and its dimension is $3n$ due to a, b, and δ in $\mathbb{F}_{2^n}$. Since the subcode contains the dual of the double error correcting BCH code, the BCH code contains the dual of our subcode. Therefore, it is clear that the minimum distance of the dual of our subcode is at least 5.

Meanwhile, we also know that if the number of nonzero distinct weights of a code is less than or equal to a minimum distance of its dual code, then its weight distribution is an explicit function of its codeword length, dimension, and distinct weights (Chapter 6, Theorem 2 in [11]). Hence, if we apply this to our subcode, we see that its weight distribution is determined by the codeword length, dimension, and distinct weights because it has at most 5 nonzero distinct weights. Consequently, the subcode has the same weight distribution as in Table II of Section II because it has the same codeword length, dimension, and weights as those of C_T. Now, we derive values of d_1 and d_2 in Lemma 9.

Lemma 10: For d_1 and d_2 in Lemma 9,

$$d_1 = \frac{1}{3}(5 \cdot 2^{n-1} + 1), \quad d_2 = \frac{1}{3}(2^{n-1} - 1).$$

Proof: For given c_i's with at least one nonzero c_i, the exponential sum of $f(x)$ has the distribution of (25). If $c_i = 0$ for $2 \le i \le l$, on the other hand, $f(x) = Tr(ax) + Tr(bx^3)$. In this case, the exponential sum of $f(x)$ has the distribution in Table I for $h = h_1 = \frac{n-1}{2}$. By summing up the distributions of both cases, the overall distribution of exponential sums of $f(x)$ is given by

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)} = \begin{cases} 2^n, & 1 \text{ time} \\ \pm 2^{n-h_1}, & (2^n-1)(d_1+1)(2^{2h_1-1} \pm 2^{h_1-1}) \text{ times} \\ 0, & (2^n-1)(d_1(2^n - 2^{2h_1}) \\ & +d_2(2^n - 2^{2h_2}) + 2^n - 2^{2h_1} + 1) \text{ times} \\ \pm 2^{n-h_2}, & (2^n-1)d_2(2^{2h_2-1} \pm 2^{h_2-1}) \text{ times} \end{cases} \tag{26}$$

for all sets of c_i's in $\mathbb{F}_{2^n}$. Moreover, this distribution is identical to the weight distribution in Table II. Note that $h_1 = \frac{n-1}{2}, h_2 = \frac{n-3}{2}$. Compared with Table II, the values of d_1 and d_2 follow immediately. ■

So far, we assume that c_i can be any element in $\mathbb{F}_{2^n}$. In sequences' aspect, however, c_i's in (24) cannot be 1 with nonzero δ. Thus, we must remove this case from the distribution in (26) in order to obtain the distribution of correlation values of sequences in $\mathcal{S}_o(2)$.

Theorem 4: The complete distribution of correlation values of any pair of binary sequences in $\mathcal{S}_o(2)$ is as follows.

$$C_{\mathbf{s}^\Lambda, \mathbf{s}^{\Lambda'}} = \begin{cases} 2^n - 1, & 2^{2n} \text{ times} \\ -1 \pm 2^{\frac{n+1}{2}}, & \frac{1}{3} \cdot 2^{2n} 2^{\frac{n-3}{2}} (2^{\frac{n-1}{2}} \pm 1) \\ & \cdot (5 \cdot 2^{2n-1} - 2^n - 5) \text{ times} \\ -1, & 2^{2n}(9 \cdot 2^{3n-4} - 3 \cdot 2^{2n-2} \\ & +3 \cdot 2^{n-2} - 1) \text{ times} \\ -1 \pm 2^{\frac{n+3}{2}}, & \frac{1}{3} \cdot 2^{2n} 2^{\frac{n-3}{2}} (2^{\frac{n-3}{2}} \pm 1) \\ & \cdot (2^{n-1} - 1)^2 \text{ times} \end{cases} \tag{27}$$

Proof: (26) shows the distribution of exponential sums of $f(x)$ for all a, b, and c_i's in $\mathbb{F}_{2^n}$. To investigate the distribution of correlation values of sequences in $\mathcal{S}_o(2)$, we need to consider the distribution of exponential sums of $f(x)$ without $c_i = 1$ ($\delta = 0$), which means removing the distribution of (25) from (26). Moreover, note that with respect to the correlation of sequences, a and b run through all elements in $\mathbb{F}_{2^n}$ by θ_0 and θ_1 as well as λ_0 and λ_1, respectively. By additionally multiplying each distribution of exponential sums by 2^{2n}, therefore, we get the distribution of correlation values of (27). ■

From Theorem 3, the linear span of sequences in $\mathcal{S}_o(2)$ has the following distribution.

$$LS(2) = \begin{cases} \frac{n(n+1)}{2}, & (2^n-1)^2 \text{ times} \\ \frac{n(n-1)}{2}, & 2(2^n-1) \text{ times} \\ \frac{n(n-3)}{2}. & 1 \text{ time} \end{cases}$$

V. An example and Implementation

In this section, we give an example of sequences in $\mathcal{S}_o(2)$ for $n = 7$ and present the implementation of sequences in $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) based on linear feedback shift registers (LFSRs).

A. An example of $\mathcal{S}_o(2)$

Consider a finite field $\mathbb{F}_{2^7}$ generated by a primitive element α satisfying $\alpha^7 + \alpha + 1 = 0$. In $\mathcal{S}_o(2)$, the kth user's sequence $\mathbf{s}^k = \{s_t^k\}$ is given by an evaluation of (17) at $x = \alpha^t, 0 \le t \le 126$, i.e.,

$$s_t^k = \begin{cases} Tr(\alpha^{i_0}\alpha^t + \alpha^{i_1}\alpha^{3t} + \alpha^{5t} + \alpha^{9t}), & i_0 \ne 127, i_1 \ne 127 \\ Tr(\alpha^{i_1}\alpha^{3t} + \alpha^{5t} + \alpha^{9t}), & i_0 = 127, i_1 \ne 127 \\ Tr(\alpha^{i_0}\alpha^t + \alpha^{5t} + \alpha^{9t}), & i_0 \ne 127, i_1 = 127 \\ Tr(\alpha^{5t} + \alpha^{9t}), & i_0 = i_1 = 127 \end{cases} \tag{28}$$

where $k = i_0 + 128 i_1$ for $0 \le i_0, i_1 \le 127$. We have 16384 cyclically distinct binary sequences in $\mathcal{S}_o(2)$. Let $\mathbf{c} = \{c_t\}$

where $c_t = Tr(\alpha^t)$, and we denote $\mathbf{c}^{(j)} = \{c_{jt}\}$, the j-decimation sequence from $\mathbf{c}$. Then, $\mathbf{c}$ is given by

$$\begin{aligned}\{c_t\} =& 10000001000001100001010001111001 \\ & 00010110011101010011111010000111 \\ & 00010010011011010110111101100011 \\ & 01001011101110011001010101111111.\end{aligned} \tag{29}$$

From (28), we see that $\{s_t^k\}$ is obtained from a linear combination of m-sequences $\mathbf{c}, \mathbf{c}^{(3)}, \mathbf{c}^{(5)}, \mathbf{c}^{(9)}$, or their shifts. In detail, for $0 \le i_0, i_1 \le 126$, $s_t^k = c_{t+i_0} + c_{3t+i_1} + c_{5t} + c_{9t}$. If either $i_0 = 127$ or $i_1 = 127$, $s_t^k = c_{3t+i_1} + c_{5t} + c_{9t}$ or $c_{t+i_0} + c_{5t} + c_{9t}$. If $i_0 = i_1 = 127$, $s_t^{16383} = c_{5t} + c_{9t}$.

From Theorem 4, the correlation distribution of sequences in $\mathcal{S}_o(2)$ for $n = 7$ is given by

$$C_{k,k'}(\tau) = \begin{cases} 127, & 16384 \text{ times} \\ 15, & 8026914816 \text{ times} \\ -17, & 6243155968 \text{ times} \\ -1, & 19127582720 \text{ times} \\ 31, & 433520640 \text{ times} \\ -33, & 260112384 \text{ times} \end{cases}$$

which is also verified from computer experiments.

B. LFSR Implementation

From the previous example, we see that the process for obtaining a sequence in $\mathcal{S}_o(\rho)$ is understood by adding $(l+1)$ m-sequences with different feedback polynomials and different initial states. Hence, $\mathcal{S}_o(\rho)$ is easy to implement by summing LFSR outputs just like Gold sequences.

For $n = 2l + 1$, $(l + 1)$ n-stage LFSRs are required to implement sequences in $\mathcal{S}_o(\rho)$ where the LFSRs have different characteristic polynomials for generating cyclically distinct m-sequences. Specifically, let $g(x)$ be a primitive polynomial over $\mathbb{F}_2$ of degree n and α be a root of $g(x)$ in $\mathbb{F}_{2^n}$. We compute $g_i(x)$ which is a minimal polynomial of α^{1+2^i} over $\mathbb{F}_2$ with $1 \le i \le l$ (for more details about computation of minimal polynomials, see [4]). Using $g_i(x)$ as the characteristic polynomial of the ith LFSR for $0 \le i \le l$ (set $g_0(x) = g(x)$), the initial states of LFSR i for $0 \le i \le \rho - 1$ can be arbitrary including zero. For $\rho \le i \le l$, on the other hand, the initial state of LFSR i is given by $\{Tr(\alpha^{(1+2^i)j})\}$ with $0 \le j \le n - 1$, which is fixed for all users. A generic description of an LFSR implementation of $\mathcal{S}_o(\rho)$ is shown in Fig. 1.

For even $n = 2l$, the sequence family $\mathcal{S}_e(\rho)$ can be implemented in the similar way to $\mathcal{S}_o(\rho)$, which is omitted here. Indeed, the implementation of $\mathcal{S}_e(\rho)$ is similar to Fig 1 except that LFSR l has a size of $\frac{n}{2}$ with a fixed initial state.

Fig. 2 shows the LFSR implementation of $\mathcal{S}_o(2)$ for $n = 7$ in the previous example. In Fig. 2, each 7-stage LFSR generates an m-sequence. The initial states of the upper two LFSRs are differently loaded according to i_0 and i_1 with a user index $k = i_0 + 128i_1$. On the other hand, the initial states of the lower two LFSRs are as shown in the figure. The initial states of LFSR 2 and LFSR 3 are given by $(c_0, c_5, c_{10}, c_{15}, c_{20}, c_{25}, c_{30}) = (1, 0, 0, 0, 0, 1, 0)$ and $(c_0, c_9, c_{18}, c_{27}, c_{36}, c_{45}, c_{54}) = (1, 0, 0, 1, 0, 1, 1)$, respectively, which can be obtained from (29). According to different initial states of upper two LFSRs, 16384 cyclically distinct binary sequences are generated to support as many different users.

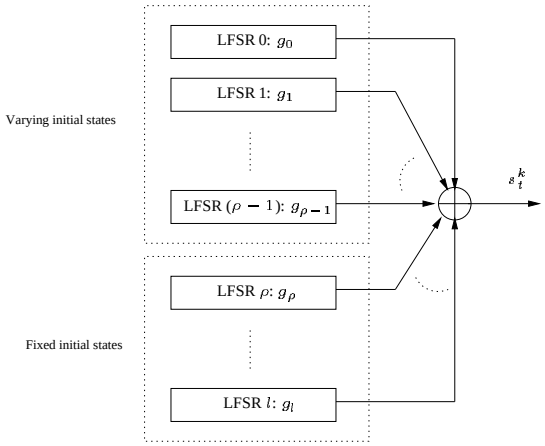

Fig. 1. LFSR implementation of $\mathcal{S}_o(\rho)$. $k = \sum_{j=0}^{\rho-1} i_j 2^{nj}$, $0 \le i_j \le 2^n - 1$. If $i_j = 2^n - 1$, then the initial state of LFSR j is set as zero.

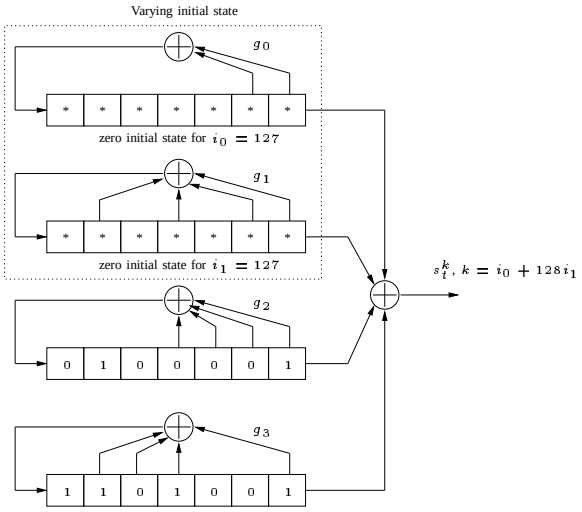

Fig. 2. LFSR implementation of $\mathcal{S}_o(2)$ for $n = 7$ with characteristic polynomials $g_0(x) = x^7 + x + 1, g_1(x) = x^7 + x^5 + x^3 + x + 1$, $g_2(x) = x^7 + x^3 + x^2 + x + 1$, and $g_3(x) = x^7 + x^5 + x^4 + x^3 + 1$.

VI. Conclusion and some observations

New families of binary sequences of period $2^n - 1$ have been presented in this paper. For odd $n = 2l + 1$ and an integer ρ with $1 \le \rho \le l$, a new family $\mathcal{S}_o(\rho)$ has family size of $2^{n\rho}$ and maximum correlation of $1 + 2^{\frac{n+2\rho-1}{2}}$. For even $n = 2l$ and an integer ρ with $1 \le \rho < l$, on the other hand, a new family $\mathcal{S}_e(\rho)$ has family size of $2^{n\rho}$ and maximum correlation of $1 + 2^{\frac{n}{2}+\rho}$. The maximum and minimum linear spans of both $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ are $\frac{n(n+1)}{2}$ and $\frac{n(n-2\rho+1)}{2}$, respectively. From the flexibility due to ρ, our new sequence families have adaptive family size and maximum correlation. The large linear spans of $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ imply that they have good potential cryptographic property.

For each ρ with $1 < \rho \leq l$, $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) contains $\mathcal{S}_o(\rho-1)$ (or $\mathcal{S}_e(\rho-1)$) as a subset. Thus, the new family contains Boztas and Kumar's case corresponding to $\mathcal{S}_o(1)$ (or Udaya's case corresponding to $\mathcal{S}_e(1)$) as a subset if m-sequences are excluded from both cases. The family $\mathcal{S}_o(2)$ is considered as a good candidate with large family size as well as low correlation. For $\mathcal{S}_o(2)$, we further derived the correlation distribution of sequences in $\mathcal{S}_o(2)$. At the end, we present the LFSR implementation of $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) and show that it is extremely easy to implement the sequence families by means of LFSR structures.

Finally, we would like to point out one interesting resemblance between our new binary sequence family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) and the quaternary sequence family $\mathcal{S}(m)$ in [8], an awarded work by Kumar, Helleseth, Calderbank, and Hammons. Jr, in 1996. From $\mathcal{S}(m)$, if we replace the trace function from Galois ring $GR(4,n)$ to Z_4 by the trace function from $\mathbb{F}_{2^n}$ to Z_2, and replace the scalar factor 2 of the sum of those monomial trace terms by a scalar factor 1, and add the sum of the remaining quadratic monomial trace terms with coefficient 1, then $\mathcal{S}(m)$ becomes $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) in our construction where $m = \rho$.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful comments. The authors' research is supported by NSERC Grant RGPIN 227700-00.

REFERENCES

[1] S. Boztas and P. V. Kumar, "Binary sequences with Gold-like correlation but larger linear span," *IEEE Trans. Inform. Theory*, vol. 40, no. 2, pp. 532-537, Mar. 1994.

[2] A. Chang, P. Gaal, S. W. Golomb, G. Gong, T. Helleseth, and P. V. Kumar, "On a conjectured ideal autocorrelation sequence and a related triple-error correcting cyclic code," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 680-687, Mar. 2000.

[3] R. Gold, "Maximal recursive sequences with 3-valued recursive cross-correlation functions," *IEEE Trans. Inform. Theory*, vol. 14, pp. 154-156, Jan. 1968.

[4] S. W. Golomb and G. Gong, *Signal Design for Good Correlation - for Wireless Communication, Cryptography and Radar*. Cambridge University Press, 2005.

[5] T. Helleseth and P. V. Kumar, *Sequences with Low Correlation*. A chapter in *Handbook of Coding Theory*. Edited by V. Pless and C. Huffman. Elsevier Science Publishers, 1998.

[6] T. Kasami, "Weight enumerators for several classes of subcodes of the 2nd order Reed-Muller codes," *Information and Control*, vol. 18, pp. 369-394, 1971.

[7] S. H. Kim and J. S. No, "New families of binary sequences with low correlation," *IEEE Trans. Inform. Theory*, vol. 49, no. 11, pp. 3059-3065, Nov. 2003.

[8] P. V. Kumar, T. Helleseth, A. R. Calderbank, and A. R. Hammons. Jr., "Large families of quaternary sequences with low correlation," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 579-592, Mar. 1996.

[9] P. V. Kumar and R. A. Scholtz, "Bounds on the linear span of bent sequences," *IEEE Trans. Inform. Theory*, vol. 29, no. 6, pp. 854-862, Nov. 1983.

[10] V. I. Levenshtein, "Bounds for codes as solutions of extremum problems for system of orthogonal polynomials," *AAECC-93, Lecture Notes in Computer Science* 673. Berlin: Springer-Verlag, pp. 25-42, 1993.

[11] F. J. MacWilliams and N. J. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland, 1977.

[12] A. A. Nechaev, "Kerdock code in a cyclic form," *Discr. Math. Appl.* vol. 1, pp. 365-384, 1991.

[13] J. S. No, S. W. Golomb, G. Gong, H. K. Lee, and P. Gaal, "Binary pseudorandom sequences of period $2^m - 1$ with ideal autocorrelation," *IEEE Trans. Inform. Theory*, vol.44, no. 2, pp. 814-817, Mar. 1998.

[14] J. D. Olsen, R. A. Scholtz, and L. R. Welch, "Bent-function sequences," *IEEE Trans. Inform. Theory*, vol. 28, no. 6, pp. 858-864, Nov. 1982.

[15] O. S. Rothaus, "Modified Gold Codes," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 654-656, Mar. 1993.

[16] A. G. Shanbhag, P. V. Kumar, and T. Helleseth, "Improved binary codes and sequence families from Z_4-linear codes," *IEEE Trans. Inform. Theory*, vol. 42, no. 5, pp. 1582-1587, Sept. 1996.

[17] V. M. Sidelnikov, "On mutual correlation of sequences," *Soviet Math. Dokl*,, vo.12, pp. 197-201, 1971.

[18] M. K. Simon, J. Omura, R. Scholtz, and K. Levitt, *Spread Spectrum Communications*, vols. I-III. Computer Science Press, Rockville, 1985.

[19] P. Udaya, "Polyphase and frequency hopping sequences obtained from finite rings," Ph.D dissertation, Dept. Elec. Eng., Indian Inst. Technol., Kanpur, 1992.

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