Design of Signal Sets with Low Intraference for CDMA Applications in Networking Environment

Guang Gong

Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1, CANADA
Email. ggong@calliope.uwaterloo.ca

Abstract. In a networking environment, interference for detecting CDMA signals is not only resulted from signals within one signal set, but also from several different signal sets when users are allowed to roam to different geographical areas. This type of interference is referred to as *intraference* among these signal sets, i.e., correlation among the signal sets. In this paper, we first provide a mathematical formalization for this concept, and exhibit the fact that constructing a family of binary signal sets with low maximum correlation is equivalent to constructing a binary signal set with larger size and low maximum correlation, which can be easily decomposed into a family of signal sets. We then derive a general formula for 0-1 distributions of the so-called one-shot sequences which has important applications in calculation of correlation functions among signal sets. Thirdly, we give constructions for three families of binary signal sets with low maximum correlation using three common signal sets: Kasami (small) signal sets (or generalized Kasami signals sets), interleaved signal sets, and bent function signal sets. We show that the family of m Kasami signal sets satisfies the mth-order shift-distinct property, which is a new concept introduced in this paper. The other interesting findings here are that (1) the maximum correlation of the enlarged interleaved signal set is comparable with the original interleaved signal set while the size is greater than the square of that of the original interleaved signal set, and (2) the maximum correlation of the enlarged bent function signal sets maintains the same maximum correlation as that of the original bent function signal set while the size is slightly smaller than the square of that of the original bent function signal set. Thus, the enlarged bent function signal set is the best among all known constructions of binary signal sets in terms of the sizes and maximum correlations.

Index Terms. Signal sets, maximum correlation, intraference, shift-distinct, CDMA.

1 Mathematical Formalization of Intraference Among Signal Sets

1.1 Introduction

In code division multiplexing access (CDMA) applications, each user is assigned a specific spreading code (or signature sequence), which is a binary sequence of period N. When multiple users transmit their signals in a common channel, the receiver (for example, the base station in mobile communication systems in the reverse link) computes the cross correlation between the received signal and each of locally generated signals and makes a decision that the one which has the largest correlation with a locally generated signal is the transmitted signal. Since the channel is shared by multiple users, the received signal is a sum of several signals. A CDMA system is designed in such a way that any pair of

signals has low cross correlation such that the receivers of different users can separate the signals that are transmitted.

More precisely, let $\mathbf{c}_j = (c_{0,j}, c_{1,j}, \cdots, c_{N-1,j}), c_{i,j} \in \{1, -1\}$ be the spreading code of user j. Let $\{a_n^{(0)}, a_n^{(1)}, \cdots, a_n^{(M-1)}\}$ be nth binary symbols that M different users designate to transmit in one symbol duration. At a receiver, the received signal is given as

$$s(t) = \sum_{n=-\infty}^{\infty} \sum_{j=0}^{M-1} a_n^{(j)} g_j(t - nT)$$

where $\{g_j(t) \,|\, j = 0, 1, \cdots, M-1\}$ is a set consisting of M pulse shapes determined by

$$g_j(t) = \sum_{i=0}^{N-1} c_{i,j} \Phi(t - iT_c)$$

where $1/T_c$ is the chip rate and $\Phi(t)$ is the chip shape waveform. After despreading and the detecting filter at the receiver of mth user, the output of the detector for detecting kth symbol $a_k^{(m)}$ is given by

$$y_k = \mu a_k^{(m)} + \sum_{n \neq k, n=-\infty}^{\infty} \sum_{j \neq m, j=0}^{M-1} a_n^{(j)} \sum_{i=0}^{N-1} c_{i,j} c_{i,m} p(kT - nT - iT_c)$$

where $p(t)$ is the convolution of $\Phi(t)$ with the receiving filter, and μ is the normalized factor. If the effect of interference resulted from $p(t)$ is negligible, then the interference to detect the symbol at the receiver of user m is determined by $\sum_{i=0}^{N-1} c_{i,j} c_{i,m}$, the correlation between $\mathbf{c}_m$ and $\mathbf{c}_j$ for each of $j \neq k, j = 0, 1, \cdots, M-1$. For more details about detection of CDMA signals, the reader is referred to [22] [21][20].

In the literature, the most of work on the signal design for spreading codes only considered the case that all users are assigned specific sequences from one particular signal set. This signal set could be, for example, a set of orthogonal sequences constructed from Hadamard matrices or binary sequences with 2-level autocorrelation, or a Gold-pair signal set, or a Kasami (small) signal set. However, in practice, since users are allowed to roam to different geographical areas, it could happen that some users' spreading codes (or sequences) may come from different signal sets. In other words, the transmitted signal not only suffers the interference from the signals in the same signal set, but also suffers the interference from the signals in different signal set. The latter is referred to as *intraference* among signals. In 2004, the author presented some preliminary results on design of signal sets with low intraference in [6]. In this paper, we will thoroughly treat this problem. In the following, we will give a mathematical definition for what we called intraference.

1.2 Intraference Among Signal Sets

A *crosscorrelation function* between two periodic binary sequences $\mathbf{a} = \{a_i\}$, of period v, and $\mathbf{b} = \{b_i\}$, of period u, over $\mathbb{F}_2$ is defined by

$$C_{\mathbf{a},\mathbf{b}}(\tau) = \sum_{i=0}^{N-1} (-1)^{a_{i+\tau}+b_i}, \tau = 0, 1, \cdots$$

where $N = \gcd(v, u)$, the greatest common factor of v and u. When $\mathbf{a} = \mathbf{b}$, the crosscorrelation between $\mathbf{a}$ and $\mathbf{b}$ becomes the autocorrelation of $\mathbf{a}$. So, sometimes, we may simply say correlation between these two sequences. For a sequence with period v, we only write the elements in one period and represent it using a vector of dimension v.

The (left) *shift operator* L performs a cyclically shifting one bit on a binary periodic sequence, i.e., $L\mathbf{a} = a_1, a_2, \cdots$ for $\mathbf{a} = a_0, a_1, \cdots$. Thus $L^r\mathbf{a} = a_r, a_{r+1}, \cdots$, is a sequence obtained by (left) shift r terms from $\mathbf{a}$ (sometimes we denote it as $L^r(\mathbf{a})$). For two sequences, if one can be obtained from the other by performing the shift operator, then we say that they are *shift equivalent*. Otherwise, they are said to be *shift distinct*.

Let $\mathbf{s}_j = (s_{j,0}, s_{j,1}, \cdots, s_{j,v-1}), 0 \leq j < r$, where each sequence cannot be obtained from other sequences by performing the shift operator. Let $S = \{\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_{r-1}\}$ and

$$\delta_S = \max |C_{\mathbf{s}_i,\mathbf{s}_j}(\tau)| \text{ for any } 0 \leq \tau < v, 0 \leq i, j < r$$

where $\tau \neq 0$ if $i = j$ (or just δ is the context is clear). The set S is said to be a (v, r, δ) *signal set*, and δ is referred to as the *maximum correlation of* S.

Two signal sets S and T are said to be *shift distinct* if each sequence in S can not be obtained from any sequence in T by performing the shift operation.

Example 1. Let S contain the following four sequences with period 15:

```
0 0 0 1 0 0 1 0 0 1 0 1 1 1 1
0 1 1 1 1 1 1 0 1 1 1 0 1 0 0
1 1 0 0 1 0 0 0 0 0 1 1 0 0 1
1 0 1 0 0 1 0 1 1 0 0 0 0 1 0
```

and T contain the following four sequences:

```
0 1 1 1 1 0 1 0 0 1 0 0 1 0 0
0 0 0 1 0 1 1 1 0 1 1 1 1 1 1
```

1 0 1 0 0 0 0 1 1 0 1 0 0 1 0

1 1 0 0 1 1 0 0 0 0 0 1 0 0 1

Then any sequence in S can not be obtained from T by the shift operation. Thus S and T are shift distinct.

Let S and T consist of shift-distinct sequences. Then S and T being shift distinct is equivalent to saying that any pair of two sequences in $S \cup T$ is shift distinct.

Definition 1. *Let S and T be two shift distinct signal sets with parameters (v, r, δ_S) and (u, s, δ_T) respectively. Let*

$$\delta_{S,T} = max|C_{\mathbf{s},\mathbf{t}}(\tau)|, \quad for\ 0 \le \tau < v, \mathbf{s} \in S, \mathbf{t} \in T.$$

Then

$$\Delta = max\{\delta_{S,T}, \delta_S, \delta_T\}$$

is said to be the maximum correlation between S and T.

For example, with the signal sets S and T in Example 1, we have $\delta_S = \max_{\mathbf{s}_i, \mathbf{s}_j \in S} |C_{\mathbf{s}_i,\mathbf{s}_j}(\tau)| = 5$, $\delta_T = 5$, and $\delta_{S,T} = 11$. Thus the maximum correlation between S and T is $\Delta = \max\{\delta_S, \delta_T, \delta_{S,T}\} = 11$.

The crosscorrelation of sequences in S represents the interference among the signals in S. On the other hand, the crosscorrelation between two sequences, each from S and T, indicates interference between two signal sets S and T. The maximum correlation between S and T is referred to as *intraference* among them in order to distinguish this type of interference from that resulted solely from the one signal set in engineering practice. The goal of design of signal sets with low intraference is to minimize the maximum correlation between signal sets S and T. In general, $\delta_{S,T}$, the maximum correlation of any pair of sequences, each from S and T, is at least the maximum value of δ_S and δ_T, the maximum correlation of the signal set S or T respectively.

Next we generalize the maximum correlation or intraference between two signal sets to correlation of a family of signal sets. In order to do so, we need to introduce a concept on mth-order shift-distinct property for signal sets.

Definition 2. *Let $S_1, \cdots, S_k$ be k signal sets in which each sequence in S_i has period v_i. We say that the signal sets $S_1, \cdots, S_k$ are mth-order shift distinct where $2 \le m \le k$ if for any signal set, its element is not a linear combination of shifts of sequences, each from any $m-1$ signal sets choosing from the remaining $k-1$ signal sets. In other words, for $\mathbf{a} \in S_i, 1 \le i \le k$,*

$$\mathbf{a} \ne c_1 L^{\tau_1} \mathbf{b}_{r_1} + \cdots + c_{m-1} L^{\tau_{m-1}} \mathbf{b}_{r_{m-1}}$$

for any $c_j \in \mathbb{F}_2$, $\tau_j \in \mathbb{Z}_{v_j}$, *and* $\mathbf{b}_{r_j} \in S_{r_j}$ *with* $r_j \neq i$, $j = 1, \cdots, m-1$.

For the family of k signal sets $S_1, \cdots, S_k$, if any two of them are shift distinct, then we say that these k signal sets are *pairwise shift distinct*, i.e., this is the case of $m = 2$. Note that the concept of k signal sets are mth-order shift distinct implies that these k signal sets are $(m-1)$th-order shift distinct. But the inverse is not true. In general, the pairwise shift-distinct property for the family of signal sets is easier to satisfy than that of the mth-order shift-distinct property for those signal sets for $m > 2$.

Definition 3. *Let* $S_1, \cdots, S_k$ *be* m*th-order* ($m \geq 2$) *shift-distinct signal sets with parameters* (v_i, r_i, δ_i), $i = 1, \cdots, k$. *We denote* δ_{S_i, S_j} *and* δ_{S_i} *by* $\delta_{i,j}$ *and* δ_i *for convenience. By this notation, we have* $\delta_i = \delta_{i,i}$. *The maximum correlation of the family of the signal sets* $S_1, \cdots, S_k$ *is defined as*

$$\Delta = \max\{\delta_{i,j}, 1 \leq i, j \leq k\}.$$

The maximum correlation of the family of the signal sets $S_1, \cdots, S_k$ indicates the maximum interference resulted from the signals in different signal sets in signal detection, which is referred to as *intraference* among these signal sets.

Design of a family of signal sets $S_1, S_2, \cdots, S_k$ with low intraference is to minimize the maximum correlation of the family. From the definition of the maximum correlation of the family of the signal sets, mathematically, it is equivalent to construct a signal set with larger size which can be decomposed into a number of signal sets with smaller sizes. We address this relation in the following subsection.

1.3 Relationship between Correlation of A Family of Signals Sets and Correlation of an Union of these Signal Sets

Let $\mathcal{S} = \cup_{j=1}^{k} S_j$ where $\{S_j \,|\, 1 \leq j \leq k\}$ are pairwise shift distinct. For any pair of two sequences $\mathbf{a}$ and $\mathbf{b}$ in $\mathcal{S}$, there are i and j such that $\mathbf{a} \in S_i$ and $\mathbf{b} \in S_j$ (here i and j may be equal). The correlation of $\mathbf{a}$ and $\mathbf{b}$, say $C_{\mathbf{a},\mathbf{b}}(\tau)$, is bounded by Δ, the maximum correlation of the family of the signal sets $S_1, \cdots, S_k$. Therefore, the maximum correlation of $\mathcal{S}$, the union of $S_j, j = 1, \cdots, k$, is equal to the maximum correlation of the family of the signal sets $S_j, j = 1, \cdots, k$. Thus, mathematically, design of the families of signal sets with low intraference is equivalent to design of the signal set $\mathcal{S}$ with low interference which is easily decomposed into a family of small signal sets , because the maximum correlation of these two cases are equal. We summarize these discussions into the following property.

Property 1. The maximum correlation of the union set $\mathcal{S} = \cup_{j=1}^{k} S_j$ is equal to the maximum correlation of the family of the signal sets $S_1, \cdots, S_k$. Furthermore, any two sequences in the union set $\mathcal{S}$ are shift distinct if and only if the family of the signal sets satisfies the pairwise shift-distinct property.

In the remaining part of this paper, we investigate design of families of signal sets with low intraference. We consider three common signal sets: Kasami (small) signal sets (including generalized Kasami signal sets), interleaved signal sets, and bent function signal sets. We found a construction for each of these three classes of signal sets with low intraference. From this approach, equivalently, we obtain three signal sets with larger sizes and low correlation, where each has parameters $(2^{2m}-1, 2^{3m}, 1+2^m)$, $((2^m-1)^2, 2^{2m}, 2^{m+1}+2^{\lfloor m/2 \rfloor+1}+1)$, and $(2^{2m}-1, 2^m(2^m-1), 1+2^m)$, respectively. Among these three new signal sets, the maximum correlation of the family of interleaved signal sets is comparable with that of an individual interleaved signal set, and the maximum correlation of the family of bent function signal sets is minimized since it is equal to the original bent function signal sets. The latter is a dramatically interesting result, since it means that we have expanded the original bent function signal set with parameters $(2^{2m}-1, 2^m, 1+2^m)$ to a new bent function signal set with parameters $(2^{2m}-1, (2^m-1)2^m, 1+2^m)$, which keeps the same maximum correlation as the original bent function signal set, but the size of the new bent function signal sets is 2^m-1 times of the original one.

The rest of the paper is organized as follow. In Section 2, we introduce some concepts and notations that will be frequently used in this paper. In Section 3, we present a general formula on 0-1 distribution of a special class of binary sequences, called one-shot sequences, which will be used to determine maximum correlation of families of Kasami signal sets and interleaved signal sets. We discuss constructions for families of Kasami signal sets, interleaved signal sets, and bent functions signal sets with low intraference in Sections 4, 5, and 6, respectively. Section 7 provides conclusions and some discussions on these signal sets.

2 Prelimilaries

The following notations and concepts will be used throughout this paper.

- The finite field $GF(2^n)$ is denoted as $\mathbb{F}_{2^n}$ for any positive integer n, and the multiplicative group of $\mathbb{F}_{2^n}$ is denoted as $\mathbb{F}_{2^n}^*$.
- The trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$ where m is a factor of n, i.e., $m|n$ is denoted as $Tr_m^n(x) = x + x^q + \cdots + x^{q^{l-1}}$ where $q = 2^m$ and $n = lm$. If the context is clear, we drop the subscript and superscript of $Tr_1^n(x)$, i.e., we write $Tr_1^n(x)$ as $Tr(x)$ for simplicity.
- α always denotes a primitive element of $\mathbb{F}_{2^n}$.
- A number x is called a *coset leader* modulo 2^n-1 if x is the smallest integer in the set $\{x, x \cdot 2 \pmod{2^n-1}, \cdots, x \cdot 2^{n-1} \pmod{2^n-1}\}$. Let Γ be the set consisting of all coset leaders modulo 2^n-1.

- Let $\{a_i\}$ be a binary sequence of period $N|2^n-1$. Using the (discrete) Fourier transform, there exists a function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ such that $a_i = f(\alpha^i), i = 0, 1, \cdots$ where $f(x) = \sum_{i\in I} Tr_1^{n_i}(\gamma_i x^i)$ where $I \subset \Gamma$, $n_i|n$ and $\gamma_i \in \mathbb{F}_{2^{n_i}}$. We say that $f(x)$ is a trace representation of $\mathbf{a}$ associated with α (for details, see [5]).

Decimation operation: A sequence with elements $a_{ri}, i = 0, 1, \cdots$, is called an r-decimation of $\mathbf{a} = \{a_i\}$, denoted as $\mathbf{a}^{(r)}$.

Interleaved Structure: A binary sequence $\mathbf{a} = \{a_i\}$ with period $N = vt$ where $1 < v, t < N$ can be arranged into an $v \times t$ array, say $A = (a_{ij})_{v\times t}$ where $a_{ij} = a_{it+j}$. This array is referred to as the $v \times t$ *array form* of $\mathbf{a}$. If all nonzero column vectors of A, considered as sequences with period v are shifts of some sequence $\mathbf{c}$, we denote the jth column of A as $A_j = (a_j, a_{t+j}, \cdots, a_{(v-1)t+j})$ (transpose symbol is omitted here), then we have $A_j = L^{e_j}(\mathbf{c}), 0 \le e_j < v$. In this case, $\mathbf{a}$ is called an *interleaved sequence*, $\mathbf{c}$ is called the *base sequence* of $\mathbf{a}$ and $(e_0, e_1, \cdots, e_{t-1})$, the *shift sequence* of $\mathbf{a}$. For details on the interleaved sequences, see [7][19][8].

Inner Product (or Dot Product): For two sequences or vectors $\mathbf{x} = (x_0, x_1, \cdots, x_{N-1})$ and $\mathbf{y} = (y_0, y_1, \cdots, y_{N-1})$, $x_i, y_i, \in \mathbb{F}_2$, we introduce the following notation.

- $(-1)^{\mathbf{x}} = ((-1)^{x_0}, (-1)^{x_1}, \cdots, (-1)^{x_{N-1}})$.
- $< \mathbf{x}, \mathbf{y} >$ denotes the inner product of $(-1)^{\mathbf{x}}$ and $(-1)^{\mathbf{y}}$, i.e., $< \mathbf{x}, \mathbf{y} >= \sum_{i=0}^{N-1}(-1)^{x_i+y_i}$.

For two constant vectors $\mathbf{0} = (0, \cdots, 0) = \mathbf{0}_N$ and $\mathbf{1} = (1, \cdots, 1) = \mathbf{1}_N$, the subscript N, the dimension, could be omitted occasionally if the context is clear.

Property 2. Let N be odd and the Hamming weight of $\mathbf{x} \neq \mathbf{0}$, denoted by $H(\mathbf{x})$, be equal to $\frac{N+1}{2}$, i.e., $H(\mathbf{x}) = \frac{N+1}{2}$. Then

1. $< \mathbf{x}, \mathbf{0} >= -1$, $< \mathbf{x}, \mathbf{1} >= 1$.
2. $< \mathbf{0}, \mathbf{0} >=< \mathbf{1}, \mathbf{1} >= N$, $< \mathbf{0}, \mathbf{1} >=< \mathbf{1}, \mathbf{0} >= -N$.

Calculation of Correlation in Terms of Exponential Sums and Inner Products: We define

$$I(\mathbf{x}) = \sum_{i=0}^{N-1}(-1)^{x_i}$$

which is referred to as the *imbalanced range* of $\mathbf{x}$. The imbalanced range of $\mathbf{x}$ gives the 0-1 distribution of the sequence $\mathbf{x}$. We list the following results which will be used throughout this paper.

1. If $N = 2^n - 1$, let $f(x)$ be a trace representation of $\mathbf{x}$, then

$$I(\mathbf{x}) + 1 = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{f(x)}. \tag{1}$$

This sum is also referred to as the exponential sum of $f(x)$.

2. For $N = vt$ with $1 < v, t < N$, let X be a $v \times t$ array form of $\mathbf{x}$ and X_j be its jth column vector, $0 \leq j < t$. Then
$$I(\mathbf{x}) = \sum_{j=0}^{t-1} I(X_j).$$
3. If $\mathbf{x}$ can be decomposed as a sum of two sequences, say $\mathbf{x} = \mathbf{a} + \mathbf{b}$, then
$$I(\mathbf{x}) = \sum_{j=0}^{t-1} < A_j, B_j >$$
where A_j and B_j are the jth column vectors of the $v \times t$ array forms of $\mathbf{a}$ and $\mathbf{b}$, respectively.
4. Let $\mathbf{a}$ and $\mathbf{b}$ be two sequences with periods $N = vt$ and $s|N$, respectively. Then the crosscorrelation between $\mathbf{a}$ and $\mathbf{b}$ can be computed in the following two different ways:
$$C_{\mathbf{a},\mathbf{b}}(\tau) = I(L^\tau \mathbf{a} + \mathbf{b}) \tag{2}$$
$$= \sum_{j=0}^{t-1} < A_j, B_j > \tag{3}$$
where A_j and B_j are jth column vectors of the array forms of τ shift of $\mathbf{a}$ and $\mathbf{b}$, respectively. Furthermore, if $N = 2^n - 1$,
$$C_{\mathbf{a},\mathbf{b}}(\tau) + 1 = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(\delta x) + g(x)}$$
where $f(x)$ and $g(x)$ are trace representations of $\mathbf{a}$ and $\mathbf{b}$ associated with α, respectively, and $\delta = \alpha^\tau$.

For a general theory on shift register sequences, the reader is referred to [4][5], detailed results on constructions of signal sets with low correlation, see [11], and deeper treatments of finite fields, see [14][15].

3 A General Formula for 0-1 Distributions of One-Shot Sequences

In this section, we formalize the sequences discussed in [9] by introducing a concept of one-shot sequences, and derive a general formula for their 0-1 distributions, which has important applications in calculation of crosscorrelation of families of Kasami (or generalized Kasami) signal sets and interleaved signal sets.

A binary sequence of period $N = vt$, $1 < v, t < N$, N odd, is called a *one-shot sequence* (with parameters (v, t)) if it can be decomposed into a sum of two sequences for which the $v \times t$ array form of one of subsequences satisfies that all the nonzero columns are balanced, i.e., the nonzero columns have the Hamming weight $\frac{v+1}{2}$, and the other subsequence has period t or it is a zero sequence. In other words, let $\mathbf{x}$ be a one-shot sequence, then we have $\mathbf{x} = \mathbf{a} + \mathbf{b}$ where $\mathbf{a} = \{a_i\}$ has period N which satisfies that the jth column vector in the $v \times t$ array form of $\mathbf{a}$, say $A = (A_0, A_1, \cdots, A_{t-1})$ where A_j's are the column

vectors of A, has the Hamming weight $\frac{v+1}{2}$ if it is not a zero vector (i.e., $H(A_j) = \frac{v+1}{2}$ if $A_j \neq \mathbf{0}$), and $\mathbf{b} = \{b_i\}$ has period t or $\mathbf{b} = \mathbf{0}$. If $\mathbf{b}$ is not a zero sequence, we may arrange $(b_0, b_1, \cdots, b_{N-1})$ into a $v \times t$ array, denoted by $B = (B_0, B_1, \cdots, B_{t-1})$ where B_j's are column vectors of B, then $B_j \in \{\mathbf{0}_v, \mathbf{1}_v\}$ because $\mathbf{b}$ has period t.

The imbalanced range of the one-shot sequences have very important applications for calculation of crosscorrelation of families of the Kasami (or generalized Kasami) signal sets and interleaved signal sets as well as other objectives in sequence design.

Theorem 1. *With the above notation, let R be the number of zero columns of A with $R \leq H(\mathbf{b})$ and $R + H(\mathbf{b}) \leq t$ where $H(\mathbf{b})$ denotes the number of 1's in one period of $\mathbf{b}$ for $\mathbf{b} \neq \mathbf{0}$, i.e., the Hamming weight of $\mathbf{b} = (b_0, b_1, \cdots, b_{t-1})$.*

Case I. *If $\mathbf{b} = \mathbf{0}$, then $I(\mathbf{x}) = I(\mathbf{a}) = R(v+1) - t$.*

Case II. *If $\mathbf{b} \neq \mathbf{0}$, then $I(\mathbf{x})$ is $(R+1)$-valued, given by*

$$I(\mathbf{x}) = \begin{cases} \pm 2j(v+1) - t + 2H(\mathbf{b}), & j = 0, 1, \cdots, k, \ R = 2k \\ \pm(2j-1)(v+1) - t + 2H(\mathbf{b}), \ j = 1, \cdots, k, & R = 2k-1. \end{cases}$$

Proof. From the discussions in Section 2, we have

$$I(\mathbf{x}) = \sum_{k=0}^{t-1} < A_k, B_k > . \tag{4}$$

Case I. $\mathbf{b} = \mathbf{0}$. In this case, (4) becomes

$$I(\mathbf{x}) = I(\mathbf{a}) = \sum_{k=0}^{t-1} < A_k, \mathbf{0} > . \tag{5}$$

Note that the case of $A_k = \mathbf{0}$ contributes v to the above sum, and $A_k \neq \mathbf{0}$ contributes -1 from Property 2-1. Thus $I(\mathbf{x}) = Rv - (t - R) = R(v+1) - t$ because there are R zero columns of A.

Case II. $\mathbf{b} \neq \mathbf{0}$. We will establish the assertion using mathematical induction. First we show that the result is true for $R = 0, 1$, and 2. Without loss of generality, we may assume that the zero vectors in $\{A_k | 0 \leq k < v\}$ (if there is any) occur at the first one or two vectors. In other words, if there is one zero vector in the column of A, we assume that $A_0 = \mathbf{0}$, and if there are two zero vectors in A, we assume that $A_0 = A_1 = \mathbf{0}$. Since B_k's are either $\mathbf{0}$ or $\mathbf{1}$ constant vectors, from Property 2, for $k \geq R$, if $B_k = \mathbf{0}$, then $< A_k, \mathbf{0} >= -1$ which contributes -1 to $I(\mathbf{x})$, if $B_k = \mathbf{1}$, $< A_k, \mathbf{1} >= 1$ contributes 1 to $I(\mathbf{x})$. We denote $r = H(\mathbf{b})$ for simplicity in the following argument.

Case 1. $R = 0$, i.e., none of A_0 and A_1 is the zero vector. In this case, there are r k's for which $< A_k, \mathbf{1} >= 1$ and, $(t - r)$ k's for which $< A_k, \mathbf{0} >= -1$. Hence, $I(\mathbf{x}) = -(t - r) + r = -t + 2r$.

Case 2. $R = 1$, i.e., $A_0 = \mathbf{0}$. In this case, by rearranging the order B_k such that $(B_{i_1}, B_{i_2}, \cdots, B_{i_{v-1}}) = (\mathbf{0}, \cdots, \mathbf{0}, \mathbf{1}, \cdots, \mathbf{1})$, we can have the following two patterns depending on $B_0 = \mathbf{0}$ or $B_0 = \mathbf{1}$ where $A_{i_k} \neq \mathbf{0}$ is denoted by $*$. (This arrangement is used in all the following cases.)

Case 2-a:

$$\begin{array}{l} A : \mathbf{0}\ * \cdots *\ * \cdots * \\ B : \mathbf{0}\ \underbrace{\mathbf{0} \cdots \mathbf{0}}_{t-r-1}\ \underbrace{\mathbf{1} \cdots \mathbf{1}}_{r} \end{array}$$

Using Property 2, $< \mathbf{0}_v, \mathbf{0}_v >= v$. This gives that $I(\mathbf{x}) = v - (t - r - 1) + r = (v + 1) - t + 2r$.

Case 2-b:

$$\begin{array}{l} A : \mathbf{0}\ * \cdots *\ * \cdots * \\ B : \mathbf{1}\ \underbrace{\mathbf{0} \cdots \mathbf{0}}_{t-r}\ \underbrace{\mathbf{1} \cdots \mathbf{1}}_{r-1} \end{array}$$

Since $< \mathbf{0}_v, \mathbf{1}_v >= -v$, this derives $I(\mathbf{x}) = -v - (t - r) + (r - 1) = -(v + 1) - t + 2r$.

Thus, if $R = 1$, then $I(\mathbf{x}) = \pm(v + 1) - t + 2r$.

Case 3. $R = 2$, i.e., $A_0 = A_1 = \mathbf{0}$.

Case 3-a: (b_0, b_1) is not a constant vector. Then we have one of B_0 and B_1 is zero. Thus $< \mathbf{0}, B_0 > + < \mathbf{0}, B_1 >= 0$ (by Property 2-2 in Section 2). In this case, $I(\mathbf{x})$ is equal to the same value as that of Case 1, i.e., $I(\mathbf{x}) = -t + 2r$. Thus, in the following, we only consider that both B_0 and B_1 are the constant zero vector or the constant one vector. Similar to Case 2, by rearranging the order of B_k, we have the following two patterns for the column vectors of A and B:

Case 3-b:

$$\begin{array}{l} A : \mathbf{00}\ * \cdots *\ * \cdots * \\ B : \mathbf{00}\ \underbrace{\mathbf{0} \cdots \mathbf{0}}_{t-r-2}\ \underbrace{\mathbf{1} \cdots \mathbf{1}}_{r} \end{array}$$

Again, using Property 2, this gives that $I(\mathbf{x}) = 2v - (t - r - 2) + r = 2(v + 1) - t + 2r$.

Case 3-c:

$$\begin{array}{l} A : \mathbf{00}\ * \cdots *\ * \cdots * \\ B : \mathbf{11}\ \underbrace{\mathbf{0} \cdots \mathbf{0}}_{t-r}\ \underbrace{\mathbf{1} \cdots \mathbf{1}}_{r-2} \end{array}$$

$\Longrightarrow I(\mathbf{x}) = -2v - (t - r) + (r - 2) = -2(v + 1) - t + 2r$. Thus, we obtain that if $R = 2$, then $I(\mathbf{x}) \in \{-t + 2r, \pm 2(v + 1) - t + 2r\}$.

Thus the result is true for $R \leq 2$. Next we assume that the result is true for $R \leq u - 1$. We will show that it is true for $R = u > 2$ and $u + r \leq t$.

Case 4-a: $(b_0, b_1, \cdots, b_{u-1})$ is not a constant vector. In this case, it is not equal to $\mathbf{0}_u$ or $\mathbf{1}_u$. Let $s_0 = |\{0 \leq j < u \,|\, b_j = 0\}|$ and $s = \min\{s_0, u - s_0\}$. For those s pairs of $< \mathbf{0}, \mathbf{0} >$ and $< \mathbf{0}, \mathbf{1} >$, they contribute zeros to $I(\mathbf{x})$. Thus, the value of $I(\mathbf{x})$ for $R = u$ is equal to the value of $I(\mathbf{x})$ computed for $R = u - 2s \geq 0$. If u is even, then $u - 2s$ is even. Since s_0 could be $1, 2, \cdots, u/2 - 1$, thus $I(\mathbf{x})$ will be

equal to $I(\mathbf{x})$ for $R = u-2, u-4, \cdots, 2, 0$. According to the induction hypothesis, we have

$$I(\mathbf{x}) = \pm 2j(v+1) - t + 2r, j = 0, 1, \cdots, u/2 - 1.$$

There are only two possible cases remained. One is $(b_0, b_1, \cdots, b_{u-1}) = \mathbf{0}_u$ and $(b_0, b_1, \cdots, b_{u-1}) = \mathbf{1}_u$ by noticing that $R = u \leq r = H(\mathbf{b})$ and $u + r \leq t$. By rearranging the order of B_k, we have the following two patterns.

Case 4-b:

$$\begin{aligned} A &: \mathbf{0}\cdots\mathbf{0}\ *\cdots*\ *\cdots* \\ B &: \underbrace{\mathbf{0}\cdots\mathbf{0}}_{u}\ \underbrace{\mathbf{0}\cdots\mathbf{0}}_{t-r-u}\ \underbrace{\mathbf{1}\cdots\mathbf{1}}_{r} \end{aligned}$$

$\Longrightarrow I(\mathbf{x}) = uv - (t-r-u) + r = u(v+1) - t + 2r.$

Case 4-c:

$$\begin{aligned} A &: \mathbf{0}\cdots\mathbf{0}\ *\cdots*\ *\cdots* \\ B &: \underbrace{\mathbf{1}\cdots\mathbf{1}}_{u}\ \underbrace{\mathbf{0}\cdots\mathbf{0}}_{t-r}\ \underbrace{\mathbf{1}\cdots\mathbf{1}}_{r-u} \end{aligned}$$

$\Longrightarrow I(\mathbf{x}) = -uv - (t-r) + (r-u) = -u(v+1) - t + 2r.$

Thus, if $(b_0, b_1, \cdots, b_{u-1})$ is a constant zero vector (Case 4-b) or one vector (Case 4-c), then $I(\mathbf{x}) = \pm u(v+1) - t + 2r$. Together with Case 4-a, we have $I(\mathbf{x}) = \pm 2j(v+1) - t + 2r, j = 0, 1, \cdots, u/2$.

If u is odd, $u - 2s$ is odd. So the similar argument for the even case can be applied to odd case. In other words, if $(b_0, b_1, \cdots, b_{u-1})$ is not of constant vectors, we can derive $I(\mathbf{x}) = \pm(2j-1)(v+1) - t + 2r, j = 1, \cdots, \frac{u-1}{2}$. If $(b_0, b_1, \cdots, b_{u-1})$ is a constant vector, then $I(\mathbf{x}) = \pm 2j(v+1) - t + 2r$ where $j = \frac{u+1}{2}$.

Thus, the result is true for $R = u$. According to mathematical induction, the result is true for any nonnegative integer $R \leq H(\mathbf{b})$ and $R + H(\mathbf{b}) \leq t$.

□

Corollary 1. *Let* $\mathbf{x}$ *be a one-shot sequence with decomposition of* $\mathbf{x} = \mathbf{a} + \mathbf{b}$ *where* $\mathbf{b}$ *has period* t *or* $\mathbf{b}$ *is a zero sequence. We assume that the number of zero columns of the array form of* $\mathbf{a}$ *is at most* R *where* $R \leq H(\mathbf{b})$ *and* $R + H(\mathbf{b}) \leq t$ *for the case* $\mathbf{b} \neq \mathbf{0}$.

1. *If* $\mathbf{b} = \mathbf{0}$, *then the imbalanced range of* $\mathbf{x}$ *is* $(R+1)$*-valued, given by*

$$I(\mathbf{x}) = j(v+1) - t, j = 0, 1, \cdots, R.$$

2. *If* $\mathbf{b} \neq \mathbf{0}$, *then the imbalanced range of* $\mathbf{x}$ *is* $(2R+1)$*-valued, given by*

$$I(\mathbf{x}) = \pm j(v+1) - t + 2H(\mathbf{b}), j = 0, 1, \cdots, R.$$

Notice that a shift of a one-shot sequence is still a one-shot sequence. Applying the identities (2) and (3) together with Corollary 1, we have the following equivalent result.

Corollary 2. *Let* $\mathbf{a}$ *be a binary sequence of period* $N = vt, 1 \leq v, t < N$ *where each column vector of the* $v \times t$ *array of* $\mathbf{a}$ *is either a zero sequence or a balanced sequence with* $(v+1)/2$ *one's, and let* $\mathbf{b}$ *be a binary sequence with period* t*. If the number of zero columns of the array form of* $\mathbf{a}$ *is at most* R *where* $R \leq H(\mathbf{b})$ *and* $R + H(\mathbf{b}) \leq t$*, then the crosscorrelation,* $C_{\mathbf{a},\mathbf{b}}(\tau)$*, between* $\mathbf{a}$ *and* $\mathbf{b}$ *is* $(2R+1)$*-valued and it belongs to*

$$C_{\mathbf{a},\mathbf{b}}(\tau) \in \{\pm j(v+1) - t + 2H(\mathbf{b}) \mid j = 0, 1, \cdots, R\}.$$

We list the values of $I(\mathbf{x})$ from Corollary 1 in Tables 1 and 2, respectively, in terms of $H(\mathbf{b})$ if $\mathbf{b} \neq \mathbf{0}$ for two particular cases. One is $\mathbf{b} = \mathbf{0}, v = 2^m - 1$, $t = 2^m + 1$, and $R = 4$, which will be used for calculation of correlation of families of Kasami signal sets and generalized kasami signal sets. The other is $\mathbf{b} \neq \mathbf{0}, v = t = 2^m - 1$ and $R = 2$, which will be applied to determine correlation of a family of interleaved signal sets.

Table 1. The Case $\mathbf{b} = \mathbf{0}$, $v = 2^m - 1$, $t = 2^m + 1$, and $R = 4$

j	$I(\mathbf{x}) = j(v+1) - t$
0	$-1 - 2^m$
1	-1
2	$-1 + 2^m$
3	$-1 + 2^{m+1}$
4	$-1 + 2^m + 2^{m+1}$

Table 2. The Case $\mathbf{b} \neq \mathbf{0}, v = t = 2^m - 1$, and $R = 2$

j	$I(\mathbf{x}) = j(v+1) - t + 2H(\mathbf{b})$
0	$-2^m + 1 + 2H(\mathbf{b})$
1	$1 + 2H(\mathbf{b})$
-1	$-2^{m+1} + 1 + 2H(\mathbf{b})$
2	$2^m + 1 + 2H(\mathbf{b})$
-2	$-2^{m+1} - 2^m + 1 + 2H(\mathbf{b})$

Remark 1. There are several examples of one-shot sequences used in calculation of correlation or Hadamard transform of sequences in the literature. For example, calculation of correlation of the Kasami (or gen-

eralized Kasami) signal sets [11] is reduced to the case $\mathbf{b} = \mathbf{0}$, $v = 2^m - 1$, $t = 2^m + 1$ and $R = 2$. For $\mathbf{b}$ not equal to a zero sequence, one example is that calculation of correlation of sequences in the interleaved signal set constructed in [7][19][8] was reduced to determining imbalanced range of the one-shot sequences for which $R = 2$, $v = t = 2^m - 1$ or a prime. The second example is that Hadamard transforms of hyper bent functions discussed in [23] were reduced to one-shot sequences with $R = 1$, $v = 2^m - 1$, $t = 2^m + 1$, the column vectors of $\mathbf{a}$ are m-sequences, and $\mathbf{b}$ is evaluated from a hyper bent function. The third example is the array correlation discussed in [9].

4 Construction of Families of Kasami (Small) and Generalized Kasami Signal Sets with Low Intraference

Let $n = 2m$, $v = 2^m - 1$, $q = 2^m$, $d = 2^m + 1$, $N = 2^n - 1$, and recall that α is a primitive element of $\mathbb{F}_{2^n}$. Let $\mathbf{s}_\lambda = \{s_{\lambda,i}\}$ be a binary sequence whose elements are given by

$$s_{\lambda,i} = f_\lambda(\alpha^i), i = 0, 1, \cdots, \text{ where} \tag{6}$$

$$f_\lambda(x) = Tr_1^m\left(Tr_m^n(x^2) + \lambda x^d\right), \lambda \in \mathbb{F}_{2^m}, x \in \mathbb{F}_{2^n}. \tag{7}$$

A signal set S consists of $\mathbf{s}_\lambda$ for all $\lambda \in \mathbb{F}_{2^m}$, i.e.,

$$S = \{\mathbf{s}_\lambda | \lambda \in \mathbb{F}_{2^m}\}. \tag{8}$$

$\mathbf{s}_\lambda$ is said to be a *Kasami (small set) sequence* and S a *Kasami (small) signal set*. Note that $f_\lambda(x)$ is the trace representation of $\mathbf{s}_\lambda$. In 1969 [10] (or [11]), Kasami established that S is a $(2^n - 1, 2^m, 2^m + 1)$ ($n = 2m$) signal set. Moreover, crosscorrelation of any pair of sequences in S or out-of-phase autocorrelation of any sequence in S is 3-valued and belongs to $\{-1, -1 \pm 2^m\}$.

For a positive integer r with $\gcd(r, 2^n - 1) = 1$, we define

$$S^{(r)} = \{\mathbf{s}^{(r)} \,|\, \mathbf{s} \in S\}.$$

In other words, each sequence in $S^{(r)}$ is the r-decimation of a sequence in S. Since the r-decimation does not change the correlation between decimated sequences when $\gcd(r, 2^n - 1) = 1$, $S^{(r)}$ is a signal set with the same parameter as those of the Kasami signal set. Furthermore, $S^{(r)}$ is shift distinct from S, which follows directly from the trace representations of the elements in S and the elements in $S^{(r)}$. So, we also call $S^{(r)}$ a Kasami signal set.

For example, according to the Kasami signal set construction, S in Example 1 can be represented as $S = \{\mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_\lambda, \mathbf{s}_{\lambda^2}\}$ where λ satisfies $\lambda^2 + \lambda + 1 = 0$, a primitive element in $\mathbb{F}_{2^2}$. This is a Kasami signal

set of $n = 4$ where α is a root of the primitive polynomial $x^4 + x + 1$ in $\mathbb{F}_{2^4}$, and $T = S^{(7)}$. For example, let $\mathbf{s}_\lambda = \{a_i\} = 110010000011001$, then $\mathbf{s}_\lambda^{(7)} = \{a_{7i}\}$, given by

$$\mathbf{s}_\lambda^{(7)} = 101000011010010$$

which is the third sequence in T.

Construction of the Family of m Kasami Signal Sets: Let $r = 2^m + 3$. A family of m Kasami signal sets is given by

$$S_i = S^{(r^i)}, i = 0, 1, \cdots, m-1.$$

For this family, we cannot establish the result on the maximum correlation of the family of m Kasami signal sets $S_0, S_1, \cdots, S_{m-1}$. But we can establish the maximum correlation between any pair of Kasami signal sets: S_{i-1} and S_i for $i = 1, \cdots, m$ where i is reduced modulo m. This means that if we place these m signal sets in a ring network topology, then we can determine maximum correlation between any two adjacent nodes which are employed signal sets S_{i-1} and S_i. In the following, first, we show that the family of these m signal sets satisfies the mth-order shift-distinct property. In order to do so, we need the following lemma about the order of r modulo $N = 2^n - 1$.

Lemma 1. *Let $r = 2^m + 3$.*

1. $\gcd(r, N) = 1$.
2. $r^m \equiv 2^m \pmod N$.

Proof. The first assertion is easy to see. We only need to establish the assertion 2. By the binomial formula,

$$r^m = (2^m + 3)^m = \sum_{i=0}^{m} \binom{m}{i} 2^{mi} 3^{m-i}. \tag{9}$$

Note that we have

$$2^{mi} \equiv \begin{cases} 2^m \pmod N \text{ for } i \text{ odd} \\ 1 \pmod N \text{ for } i \text{ even.} \end{cases}$$

Thus, we can rewrite (9) as follows:

$$r^m \equiv L_0 + 2^m L_1 \pmod N \tag{10}$$

where

$$L_k = \sum_{j=0}^{\lfloor m/2 \rfloor - k} \binom{m}{2j+k} 3^{m-(2j+k)}, \ k = 0, 1.$$

On the other hand, we have

$$2^{2m} = (3+1)^m = \sum_{i=0}^{m} \binom{m}{i} 3^i = L_0 + L_1 \tag{11}$$

$$2^m = (3-1)^m = \sum_{i=0}^{m} \binom{m}{i} (-1)^i 3^i = L_0 - L_1. \tag{12}$$

Thus

$$\begin{aligned} (11) + (12) \Longrightarrow \quad & 2^{2m} + 2^m = 2L_0 \\ \Longrightarrow\ & L_0 = 2^{2m-1} + 2^{m-1}. \end{aligned}$$

Substituting this value to (12),

$$L_1 = L_0 - 2^m = 2^{2m-1} + 2^{m-1} - 2^m = 2^{2m-1} - 2^{m-1}.$$

Substituting the values of both L_0 and L_1 into (10), we get

$$r^m \equiv 2^{2m-1} + 2^{m-1} + 2^m(2^{2m-1} - 2^{m-1}) \equiv 2^m \pmod{N}.$$

□

We need the following result to determine two functions being equal. (The proof of the following lemma can be found in a finite field book, say [14][15] or in [5].)

Lemma 2. *Let $h(x) = \sum_{i=0}^{2^n-1} h_i x^i$ and $g(x) = \sum_{i=0}^{2^n-1} g_i x^i$ be two functions from $\mathbb{F}_{2^N}$ to $\mathbb{F}_2$, $h_i, g_i \in \mathbb{F}_{2^n}$, then $h(x) = g(x)$ if and only if $h_i = g_i, 0 \le i < 2^n$. Furthermore, if we write both $h(x)$ and $g(x)$ in their trace representations, i.e., $h(x) = \sum_{i\in\Gamma} Tr_1^{n_i}(h_i x^i)$ and $g(x) = \sum_{i\in\Gamma} Tr_1^{n_i}(g_i x^i)$ where Γ is the set consisting of all coset leaders modulo $2^n - 1$, $n_i | n$, and $h_i, g_i \in \mathbb{F}_{2^{n_i}}$, then $h(x) = g(x)$ if and only if $h_i = g_i, \forall i \in \Gamma$, and $h(0) = g(0)$.*

Theorem 2. *$S_i, 0 \le i < m$, m Kasami signal sets, are mth-order shift distinct.*

Proof. Note that we may write (7) as

$$f_\lambda(x) = Tr_1^n(x^2) + Tr_1^m(\lambda x^d) = Tr_1^n(x) + Tr_1^m(\lambda x^d)$$

because $Tr_1^n(x^2) = Tr_1^n(x)$ in the construction of the Kasami signal set (the square on x is needed for the proof of the correlation property). Therefore, we have $f_\lambda(x^{r^i}) = Tr_1^n(x^{r^i}) + Tr_1^m(\lambda x^{r^i d})$, which is the trace representation of a sequence in S_i. First, we show that for any $i : 0 < i < m$, S_i and S_0 are shift distinct which implies that any pair of signal sets are shift distinct. Let $\mathbf{a} \in S_0$ and $\mathbf{b} \in S_i$ with trace representations $f_\lambda(x)$ and $f_\theta(x^{r^i})$, respectively. If $\mathbf{a}$ is a shift of $\mathbf{b}$ by τ, then we have

$$Tr_1^n(x) + Tr_1^m(\lambda x^d) = Tr_1^n((\delta x)^{r^i}) + Tr_1^m(\theta(\delta x)^{2^i d}),\ \delta = \alpha^\tau \tag{13}$$

the last term comes from the fact that $r^i d \equiv 2^i d \bmod (2^n - 1)$. According to Lemma 1, the exponents in the monomial trace terms of $Tr_1^n(\cdot)$ in (13) are not in the same coset modulo $2^n - 1$. Applying Lemma 2, (13) is not valid. So, S_0 and S_i are shift distinct for $i = 1, \cdots, m-1$. Using a similar argument to the above case, we can show that $f_\lambda(x)$ cannot be written as a linear combination of shifts of the sequences taking from more than two signal sets. In other words, there exist no $c_i \in \mathbb{F}_2$ and $\delta_i \in \mathbb{F}_{2^n}^*$ such that $f_\lambda(x) = \sum_{i=1}^{m-1} c_i f_\lambda(\delta_i x^{r^i})$. If so, we have

$$Tr_1^n(x) + Tr_1^m(\lambda x^d) = \sum_{i=1}^{m-1} c_i [Tr_1^n(\delta_i x^{r^i}) + Tr_1^m(\lambda \delta_i x^{dr^i})]. \tag{14}$$

However, the right-hand side of (14) has at least $k + 1 > 2$ distinct monomial trace terms where $k > 1$ and k is the number of nonzero coefficients c_i while the left-hand side of (14) has only two distinct monomial trace terms. According Lemma 2, (14) is not valid. Thus a sequence in S_0 cannot be equal to a linear combination of shifts of sequences in $S_i, i = 1, \cdots, m-1$. Thus, $S_i, 0 \le i < m$ are mth-order shift distinct.

□

Theorem 3. *For a fixed $i, 1 \le i \le m$ where i is reduced modulo m, the maximum crosscorrelation of two signal sets S_{i-1} and S_i, denoted as Δ, is given by*

$$\Delta = 2^{m+1} + 2^m - 1.$$

Moreover, correlation of any pair of sequences $\mathbf{a}$ and $\mathbf{b}$ where $\mathbf{a} \in S_{i-1}$ and $\mathbf{b} \in S_i$ or both $\mathbf{a}$ and $\mathbf{b}$ taking from the same signal set or the out-of-phase autocorrelation of the sequences in the family is five-valued and belongs to $\{-1 + c2^m \mid c = 0, \pm 1, 2, 3\}$.

A proof given here is similar to the proof of the crosscorrelation of the sequences in the Kasami signal set. To prove Theorem 3, we need the following lemmas. The following assertions related to finite fields and trace functions which will be used in several places.

Assertions:

1. For any $y \in \mathbb{F}_{2^m}, x \in \mathbb{F}_{2^n}$, $Tr_m^n(xy) = yTr_m^n(x)$.
2. For any $x, y \in \mathbb{F}_{2^m}$, $Tr_1^m(x^{2^t} y) = Tr_1^m(xy^{2^{m-t}})$.

Lemma 3. *Let*

$$f_c(x) = Tr_1^n(x^2) + Tr_1^m(cx^d), c \in \mathbb{F}_q,\ x \in \mathbb{F}_{2^n}$$

which is the trace representation of a sequence $\mathbf{s}_c$ in the Kasami signal set S_0. For $X = zx, z \in \mathbb{F}_q^*, x \in \mathbb{F}_{q^2}$ (recall $q = 2^m$), $f_c(x)$ can be represented as

$$f_c(zx) = Tr_1^m(z^2 t(x, c)) \tag{15}$$

where

$$t(x,c) = Tr_m^n(x^2) + cx^d. \tag{16}$$

Proof. For $X = zx, z \in \mathbb{F}_q^*$,

$$\begin{aligned} f_c(X) &= Tr_1^n(z^2x^2) + Tr_1^m(cz^dx^d) \\ &= Tr_1^m(z^2Tr_m^n(x^2)) + Tr_1^m(cz^dx^d) \\ &= Tr_1^m(z^2[Tr_m^n(x^2) + cx^d]) \end{aligned}$$

where the last identity comes from $z^d = z^2, z \in \mathbb{F}_q^*$. Thus, the assertion is true.

□

Lemma 4.

$$\begin{aligned} t^2(x,c) &= x^4(1 + y^4 + c^2y^2) \\ t^{2^{m-1}}(x^r,c) &= x^4(y^3 + y + c^{2^{m-1}}y^2) \end{aligned}$$

where $y = x^{2^m-1}$.

Proof. Note that

$$\begin{aligned} t^2(x,c) &= (Tr_m^n(x^2) + cx^d)^2 \\ &= (x^2 + x^{2\cdot 2^m} + cx^{2^m+1})^2 \\ &= x^4 + x^{2^{m+2}} + c^2x^{2d} \\ &= x^4(1 + x^{4(2^m-1)} + c^2x^{2(2^m-1)}) \end{aligned}$$

$$\implies t^2(x,c) = x^4(1 + y^4 + c^2y^2), \quad \text{where } y = x^{2^m-1}. \tag{17}$$

We apply a similar process as that of $t^2(x,c)$ to $t^{2^{m-1}}(x^r,c)$.

$$\begin{aligned} t^{2^{m-1}}(x^r,c) &= (Tr_m^n(x^{2r}) + cx^{4d})^{2^{m-1}} \\ &= x^{2^mr} + x^r + c^{2^{m-1}}x^{2d} \\ &= x^{3\cdot 2^m+1} + x^{2^m+3} + c^{2^{m-1}}x^{2^{m+1}+2} \\ &= x^4(x^{3(2^m-1)} + x^{2^m-1} + c^{2^{m-1}}x^{2(2^m-1)}) \\ &= x^4(y^3 + y + c^{2^{m-1}}y^2) \end{aligned}$$

That's,

$$t^{2^{m-1}}(x^r,c) = x^4(y^3 + y + c^{2^{m-1}}y^2). \tag{18}$$

□

Proof of Theorem 3. It is enough to show the result is true for $i = 1$. To avoid the double indexes, we write $\mathbf{a} = \mathbf{s}_\lambda \in S_0$ and $\mathbf{b} = \mathbf{s}_\theta^{(r)} \in S_1$. Set $\delta = \alpha^\tau$. Applying Lemma 3 to $f_\lambda(X\delta)$ and $f_\theta(X^r)$, the trace representations of $L^\tau\mathbf{a}$ and $\mathbf{b}$, for $X = zx, z \in \mathbb{F}_q^*$, we get

$$\begin{aligned} f_\lambda(X\delta) &= f_\lambda(z(x\delta)) = Tr_1^m(z^2 t(x\delta, \lambda)) \\ &= Tr_1^m(z^4 t^2(x\delta, \lambda)) \qquad (19) \\ f_\theta(X^r) &= f_\theta(z^r x^r) = Tr_1^m(z^{2r} t(x^r, \theta)) \\ &= Tr_1^m(z^8 t(x^r, \theta)) \ (z^r = z^4, \text{ for } z \in \mathbb{F}_q^*) \\ &= Tr_1^m(z^4 t^{2^{m-1}}(x^r, \theta)) \text{ (by Assertion 2)} \qquad (20) \end{aligned}$$

Let A and B be the $v \times d$ arrays from $L^\tau(\mathbf{a})$ and $\mathbf{b}$, respectively, and let A_j and B_j, $0 \le j < d$ be column vectors of A and B, respectively. From (19) and (20), we get the trace representations of A_j and B_j, say $f_{A_j}(z)$ and $f_{B_j}(z)$, given by

$$f_{A_j}(z) = Tr_1^m(z^4 t^2(\alpha^j \delta, \lambda)) \qquad (21)$$

$$f_{B_j}(z) = Tr_1^m(z^4 t^{2^{m-1}}(\alpha^{jr}, \theta)) \qquad (22)$$

where $t(x, y)$ is defined by (16) in Lemma 3. Thus $C_{\mathbf{a},\mathbf{b}}(\tau)$, the crosscorrelation of $\mathbf{a}$ and $\mathbf{b}$, is equal to the sum of the inner products of $(-1)^{A_j}$ and $(-1)^{B_j}$, $0 \le j < d$, or equivalently, the imbalanced range of $A_j + B_j$, i.e.,

$$C_{\mathbf{a},\mathbf{b}}(\tau) = \sum_{j=0}^{d-1} < A_j, B_j > = \sum_{j=0}^{d-1} I(A_j + B_j).$$

Note that we can consider $L^\tau(\mathbf{a}) + \mathbf{b}$ as a one-shot sequence with parameters $v = 2^m - 1$ and $d = 2^m + 1$, where the shorter sequence is a zero sequence. Thus, it suffices to prove that there are at most four identical corresponding columns in the matrices A and B. This is equivalent to saying that there are at most four values of j among $0 \le j < d$ such that

$$t^2(\alpha^j \delta, \lambda) = t^{2^{m-1}}(\alpha^{jr}, \theta) \qquad (23)$$

where $t(x, y)$ is defined by (16). From Lemma 4, the above equation becomes

$$t^2(x\alpha^\tau, s) = x^4(\alpha^{4\tau} + \alpha^{\tau 2^{m+2}} y^4 + \lambda^2 \alpha^{\tau 2d} y) = x^4(y^3 + y + \theta^{2^{m-1}} y^2) = t^{2^{m-1}}(x^r, \theta).$$

Simplifying that, we obtain a quartic equation

$$ay^4 + by^3 + cy^2 + dy + e = 0 \qquad (24)$$

where

$$a = \alpha^{\tau 2^{m+2}}, b = 1, c = \theta^{2^{m-1}}, d = \lambda^2 \alpha^{\tau 2d} + 1, e = \alpha^{4\tau}$$

where $\tau \neq 0$. Since (24) is a quartic equation in $\mathbb{F}_{2^n}$, it has at most four solutions $y_i, i = 1, 2, 3, 4$ in $\mathbb{F}_{2^n}$. For each of the y_i, there exists at most one $x = \alpha^j, 0 \leq j < d$ such that

$$y_i = x^{2^m - 1}.$$

Therefore, there are at most four values of j among $0 \leq j < d$ such that (23) is true. Applying Table 1 in Section 3 for the imbalanced range of $L^\tau(\mathbf{a}) + \mathbf{b}$, considered as a one-shot sequence with $R = 4$, together with the crosscorrelation of sequences in the individual signal set S_0 or S_1, the result follows.

□

From Theorem 2, the proof of Theorem 3 and Property 1, we have the following corollary.

Corollary 3. *Let $g_{b,c}(x) = Tr_1^n(x + bx^r) + Tr_1^m(cx^d), b \in \mathbb{F}_{2^n}, c \in \mathbb{F}_{2^m}$. Let a signal set K^+ consist of all the sequences whose trace representations are $g_{b,c}(x)$ when b runs through $\mathbb{F}_{2^n}$ and c runs through $\mathbb{F}_{2^m}$. Then K^+ is a signal set with parameters $(2^{2m} - 1, 2^{3m}, 3 \cdot 2^m - 1)$. Furthermore, the crosscorrelation of any two sequences in K^+ or out of phase autocorrelation of any sequence in K^+ is five-valued and belongs to $\{-1 + c2^m \mid c = 0, \pm 1, 2, 3\}$.*

From Corollary 3, we have the following equivalent result in the coding context.

Corollary 4. *Let $n = 2m$, $r = 2^m + 3$, and*

$$f(x) = Tr_1^n(ax + bx^r) + Tr_1^m(cx^{r-2}), a, b \in \mathbb{F}_{2^n}, c \in \mathbb{F}_{2^m}.$$

Let α be a primitive element of $\mathbb{F}_{2^n}$, $s_i = f(\alpha^i), i = 0, 1, \cdots, 2^n - 2$, and $\mathcal{C}$ contains all $\mathbf{s} = \{s_i\}$ when a, b run through $\mathbb{F}_{2^n}$ and c runs through $\mathbb{F}_{2^m}$. Then $\mathcal{C}$ is a cyclic BCH code with the minimum distance $2^{2m-1} - 2^m - 2^{m-1}$ and the Hamming weight of any codeword in $\mathcal{C}$ belongs to the set $\{2^{2m-1} - c2^{m-1} \mid c = 0, \pm 1, 2, 3\}$.

Remark 2. Cusick and Dobbertin (1996, [1]) established that a subcode of $\mathcal{C}$, given by $g(x) = Tr_1^n(ax + bx^r), a, b \in \mathbb{F}_{2^n}$, has three weights of $2^{2m-1}, 2^{2m-1} \pm 2^{m-1}$. Recently, Dobbertin *et al.* [2] established that a subset of $\mathcal{C}$ are bent functions.

A function g from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$ is said to be *orthogonal* if it is a trace representation of some binary 2-level autocorrelation sequence, i.e., if $\mathbf{b} = \{b_i\}$ where $b_i = g(\eta^i), i = 0, 1, \cdots$ has 2-level autocorrelation where η is a primitive element of $\mathbb{F}_{2^m}$, then we say that g is orthogonal. See [5] for details about the orthogonal functions defined in terms of 2-level autocorrelation sequences. The trace function $Tr_1^m(x)$

employed in the Kasami (small) set construction can be replaced by any orthogonal function from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$, which will be discussed below.

A family of m generalized Kasami signal sets is obtained from the Kasami signal set, defined in (8), in which $Tr_1^m(x)$ is replaced by an arbitrary orthogonal function g from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$. In other words, the function $f_\lambda(x)$ defined in (7) is replaced by

$$f_\lambda(x) = g\left(Tr_m^n(x^2) + \lambda x^d\right), \lambda \in \mathbb{F}_{2^m}, x \in \mathbb{F}_{2^n}. \tag{25}$$

The generalized Kasami signal set, denoted as $S(g)$, is a $(2^n - 1, 2^m, 2^m + 1)$ $(n = 2m)$ signal set for any orthogonal function g. Furthermore, the crosscorrelation of any pair of sequences in $S(g)$ or out-of-phase autocorrelation of a sequence in $S(g)$ is 3-valued and belongs to $\{-1, -1 \pm 2^m\}$.

This result was first established by No and Kumar in 1989 [17] for $g(x) = Tr_1^m(x^s)$ where $\gcd(s, 2^m - 1) = 1$, and later No, Yang, Chung, and Song [16] showed this result for some particular orthogonal functions in 1997. The result being valid for an arbitrary orthogonal function g was observed by the author in 2002 [8].

We denote this family of m generalized Kasami signal sets by

$$T_i = S(g)^{(r^i)}, i = 0, 1, \cdots, m - 1.$$

The following result on the maximum correlation for m generalized Kasami signal sets is the same as that for the m Kasami signal sets $S_i : 0 \le i < m$.

Theorem 4. *For a fixed $i, 1 \le i \le m$ (i is reduced modulo m), the maximum crosscorrelation of two signal sets T_{i-1} and T_i, denoted as Δ, is given by*

$$\Delta = 2^{m+1} + 2^m - 1.$$

Moreover, correlation of any pair of sequences $\mathbf{a}$ and $\mathbf{b}$ where $\mathbf{a} \in T_{i-1}$ and $\mathbf{b} \in T_i$ or both $\mathbf{a}$ and $\mathbf{b}$ taking from the same signal set or the out-of-phase autocorrelation of sequences in one of these sets is five-valued and belongs to $\{-1 + c2^m \mid c = 0, \pm 1, 2, 3\}$.

Proof. The proof can be proceeded in the same way as those for Theorem 3 until the step that we derive trace representations of the column sequences A_j and B_j of $L^\tau\mathbf{a}$ and $\mathbf{b}$ ($\mathbf{a} \in T_{i-1}$ and $\mathbf{b} \in T_i$), respectively, which are replaced by:

$$f_\lambda(X\delta) = f_\lambda(zx\delta) = g(z^4 t^2(x\delta, \lambda)) \tag{26}$$

$$f_\theta(X^r) = f_\theta(z^r x^r) = g(z^4 t^{2^{m-1}}(x^r, \theta)) \tag{27}$$

where $X = zx, z \in \mathbb{F}_q^*$. Hence A_j and B_j are 2-level autocorrelation sequences given by $g(x)$ with the same shifts or distinct shifts. If they have the same shift, then it contributes v to the crosscorrelation function between $\mathbf{a}$ and $\mathbf{b}$, otherwise it contributes -1 to the crosscorrelation function between $\mathbf{a}$ and $\mathbf{b}$. Since the phase values in the generalized Kasami case is the same as the Kasami case from equations (26) and (27), continuing the same process as in the proof of Theorem 3, the result follows.

□

Note. It is not easy to show whether m generalized Kasami signal sets are kth-order shift distinct for $2 < k \leq m$, when $g(x)$ is not equal to $Tr_1^m(x)$. Showing the kth-order shift-distinct property for $T_i, 0 \leq i < m$ needs to use the result of the expansion of $f_\lambda(x^{r^i}) = g(Tr_m^n(x^{r^i}) + \lambda x^{r^i d})$ which is unknown, because the number of terms in this expansion is equal to the linear span of the sequences in T_i. So, it seems difficult to establish the kth-order shift-distinct property for $2 < k \leq m$ even for a simple case of g. For the pairwise shift-distinct property of the family of those m signal sets, we believe that the result is true. However, we still work on that for getting a nice proof.

5 Construction of A Family of Interleaved Signal Sets with Low Intraference

In order to construct a family of $v + 2$ interleaved signal sets, we need two ingredients: the interleaved signal sets and the Gold-pair signal sets.

Interleaved Signal Sets: A binary interleaved signal set is a signal set with parameters $(v^2, v+1, 2v+3)$ where $v = 2^m - 1$ or v is a prime, which can be constructed as follows. Let $\mathbf{u} = (u_0, u_1, \cdots, u_{v^2-1})$ be a (v, v) interleaved sequence whose jth column sequence in the $v \times v$ array form is given by $L^{e_j}(\mathbf{a})$, i.e., $u_{iv+j} = a_{i+e_j}$ where $\mathbf{a} = (a_0, a_1, \cdots, a_{v-1})$ is a binary sequence of period v with 2-level autocorrelation, and $\mathbf{e} = (e_0, e_1, \cdots, e_{v-1})$ is an integer sequence whose elements are taken from $\mathbf{Z}_v$, which satisfies the following difference condition:

$$\begin{gathered}\text{for each } 1 \leq s < v, \text{ the differences} \\ e_j - e_{j+s},\ 0 \leq j < v - s \text{ are all distinct.}\end{gathered}$$

Let

$$\mathbf{s}_j = (s_{j,0}, s_{j,1}, \cdots, s_{j,v^2-1}),\ 0 \leq j < v$$

whose elements are defined by

$$s_{j,i} = u_i + b_{j+i} \text{ or } \mathbf{s}_j = \mathbf{u} + L^j(\mathbf{b}), 0 \leq j < v$$

where $\mathbf{b}$ is a binary sequence of period v with 2-level autocorrelation. A signal set $S(\mathbf{b})$ defined as

$$S(\mathbf{b}) = \{\mathbf{s}_j \mid j = 0, 1, \cdots, v-1\} \cup \{\mathbf{u}\}.$$

is an interleaved signal set, which has parameters $(v^2, v+1, 2v+3)$. Moreover, the crosscorrelation of any two sequences in $S(\mathbf{b})$ or the out-of-phase auto-correlation of any sequence in $S(\mathbf{b})$ belongs to the set $\{1, -v, v+2, 2v+3, -2v-1\}$. Especially, the autocorrelation of the interleaved sequence $\mathbf{u}$ is three-valued and belongs to $\{1, -v, v+2\}$. (See Gong and Paterson's work for the properties of interleaved sequences in [7][19][8].)

Gold-Pair Signal Sets: Let $\mathbf{b}$ be an m-sequence of degree m, and let $\mathbf{c}$ be a d-decimation of $\mathbf{b}$, i.e., $c_i = b_{di}, i = 0, 1, \cdots$. A Gold-pair signal set S is given by $S = \{\mathbf{b} + L^i\mathbf{c} \,|\, 0 \le i < 2^m - 1\} \cup \{\mathbf{b}, \mathbf{c}\}$ where d is chosen from the following known cases.

For m odd,

(a) $d = 2^k + 1$ (the Gold decimation), $\gcd(k, m) = 1$ and $k \le \frac{m-1}{2}$.

(b) $d = 2^{2k} - 2^k + 1$ (the Kasami (large set) decimation), $\gcd(k, m) = 1$ and $k \le \frac{m-1}{2}$.

(c) $d = 2^{\frac{m-1}{2}} + 3$ (the Welch decimation).

(d) $d = 2^{2k} + 2^k - 1$ (the Niho decimation) where

$$k = \begin{cases} \frac{m-1}{4} & \text{if } m \equiv 1 \pmod 4 \\ \frac{3m-1}{4} & \text{if } m \equiv 3 \pmod 4 \end{cases}.$$

(e) Inverse of d, for d in each of the above four cases.

For m even,

(f) $m = 2r$ where r is odd, the known cases for d are: $d = 2^{r+1} + 2^{(r+1)/2} + 1$ or $d = 2^{r+1} + 3$ and their inverses (Cusik, Dobbertin, 1996 [1]).

Then S is a Gold-pair signal set with parameters $(2^m - 1, 2^m + 1, 1 + 2^{\lfloor m/2 \rfloor + 1})$. Furthermore, the correlation of any pair of sequences in S takes three values of $-1, \pm 2^{\lfloor m/2 \rfloor + 1}$. Equivalently, the Hamming weight of any sequence with trace representation $Tr(\lambda x + \mu x^d), \lambda, \mu \in \mathbb{F}_{2^m}$ belongs to $\{2^{m-1}, 2^{m-1} \pm 2^{\lfloor m/2 \rfloor + 1}\}$. (See details about the Gold-pair sequences in the original Gold's paper in 1969 [3] or Helleseth and Kumar's excellent chapter [11].)

Construction of A Family of $v + 2$ Interleaved Signal Sets for $v = 2^m - 1$:

Let $\mathbf{b}$ be an m-sequence of degree m, and $\mathbf{c}$ be a d-decimation of $\mathbf{b}$, i.e., $c_i = b_{id}, i = 0, 1, \cdots$. We assume that d is taken from the above list which gives the Gold-pair sequences. We construct $v + 2$ interleaved signal sets with parameters $(v^2, v + 1, 2v + 3)$ where $v = 2^m - 1$ as follows. Denoting $\mathbf{b}_{i+1} = \mathbf{b} + L^i(\mathbf{c}), i = 0, \cdots, 2^m - 2$, $\mathbf{b}_0 = \mathbf{b}$, and $\mathbf{b}_{2^m} = \mathbf{c}$ for convenience, i.e., $\{\mathbf{b}_i \,|\, 0 \le i \le 2^m\}$ constitutes a Gold signal set. We define

$$IS_i = S(\mathbf{b}_i) \setminus \{\mathbf{u}\}, 0 \le i < v + 2.$$

The family $\{IS_i\}$ contains the $v+2 = 2^m+1$ interleaved signal sets, each with parameters $((2^m-1)^2, 2^m-1, 2^{m+1}+1)$ according to the construction of interleaved signal sets. Here the sequence $\mathbf{u}$ is excluded in order to satisfy the pairwise shift-distinct property for the family.

Theorem 5. *The 2^m+1 interleaved signal sets $IS_i, 0 \le i < 2^m$ with parameters $((2^m-1)^2, 2^m-1, 2^{m+1}+1)$ are pairwise shift distinct.*

Proof. We first consider $\mathbf{s} \in IS_i$ and $\mathbf{t} \in IS_j$ with $0 < i \neq j < 2^m$. Then we have

$$\mathbf{s} = \mathbf{u} + L^k\mathbf{b} + L^{k+i}\mathbf{c}$$
$$\mathbf{t} = \mathbf{u} + L^l\mathbf{b} + L^{l+j}\mathbf{c}$$

If they are shift equivalent, then we get

$$\mathbf{s} = L^\tau\mathbf{t} \Longrightarrow \mathbf{u} + L^k\mathbf{b} + L^{k+i}\mathbf{c} = L^\tau\mathbf{u} + L^{l+\tau}\mathbf{b} + L^{l+j+\tau}\mathbf{c}. \tag{28}$$

We may write $\mathbf{u} = \mathbf{x}+\mathbf{y}$ where $\mathbf{y}$ has period v ($v = 2^m-1$) and $\mathbf{x}$ has period v^2 for which the monomial trace terms in its trace representation has degree that is not a divisor of m. Note that both $\mathbf{b}$ and $\mathbf{c}$ are m-sequences of degree m with period v, and $L^\tau\mathbf{u} = L^\tau\mathbf{x} + L^\tau\mathbf{y}$. From Lemma 2 and the trace representation of $\mathbf{x}$, the identity (28) forces $\mathbf{x} = L^\tau\mathbf{x} \Longrightarrow \tau = 1$. Substituting it to (28), we get

$$L^k\mathbf{b} + L^{k+i}\mathbf{c} = L^{l+\tau}\mathbf{b} + L^{l+j+\tau}\mathbf{c}. \tag{29}$$

Since $\mathbf{c}$ is a d-decimation of $\mathbf{b}$, again using Lemma 2, the identity (29) is not valid.

If the case that both $\mathbf{s}$ and $\mathbf{t}$ belong to IS_0 or IS_{2^m} occurs or the case that one belongs to IS_0 and the other belongs to IS_{2^m} occurs, then we still get an identity similar to (29), but one term of shift of $\mathbf{b}$ or one term of shift of $\mathbf{c}$ does not occur. However, it is still not valid according to Lemma 2. So, $\mathbf{s}$ and $\mathbf{t}$ are not shift equivalent.

□

Theorem 6. *The maximum correlation of the family of 2^m+1 interleaved signal sets $IS_i, 0 \le i \le 2^m$, Δ, is given by*

$$\Delta = 2^{m+1} + 2^{\lfloor m/2 \rfloor + 1} + 1.$$

In other words, for any pair of sequences $\mathbf{s} \in IS_i$ and $\mathbf{t} \in IS_j$, the crosscorrelation between $\mathbf{s}$ and $\mathbf{t}$ is bounded by

$$|C_{\mathbf{s},\mathbf{t}}(\tau)| \le 2^{m+1} + 2^{\lfloor m/2 \rfloor + 1} + 1, 0 \le \tau < (2^m-1)^2, \tau \neq 0 \text{ if } \mathbf{s} = \mathbf{t}.$$

Proof. We will use Corollary 1 in Section 3 to establish the result. For $\mathbf{s} \in IS_i$ and $\mathbf{t} \in IS_j$, we have

$$\mathbf{s} = \mathbf{u} + L^k(\mathbf{b}_i) \tag{30}$$

$$\mathbf{t} = \mathbf{u} + L^l(\mathbf{b}_j). \tag{31}$$

The crosscorrelation between $\mathbf{s}$ and $\mathbf{t}$ depends on the Hamming weight of

$$L^\tau \mathbf{s} + \mathbf{t} = (\mathbf{u} + L^\tau \mathbf{u}) + (L^{k+\tau}\mathbf{b}_i + L^l \mathbf{b}_j). \tag{32}$$

Let $X = [X_0, X_1, \cdots, X_{v-1}]$ and $W = [W_0, W_1, \cdots, W_{v-1}]$ be the $v \times v$ array forms of $\mathbf{x} = \mathbf{u} + L^\tau \mathbf{u}$ and $\mathbf{w} = L^{k+\tau}\mathbf{b}_i + L^l \mathbf{b}_j$, respectively, where X_j's and W_j's are column vectors of X and W respectively. Then

$$C_{\mathbf{s},\mathbf{t}}(\tau) = \sum_{j=0}^{v-1} < X_j, W_j > .$$

Let $\mathbf{y} = \mathbf{x} + \mathbf{w}$. We will show that $\mathbf{y}$ is a one-shot sequence. Then we can apply Corollary 1 in Section 3 to determine their correlation values.

If $\mathbf{w} = \mathbf{0}$, then $\mathbf{y} = \mathbf{x} + \mathbf{w} = \mathbf{u} + L^\tau \mathbf{u}$ where $\mathbf{u}$ is the interleaved sequence defined in the construction of the interleaved signal set. Thus

$$C_{\mathbf{s},\mathbf{t}}(\tau) = \sum_{j=0}^{v-1} < X_j, \mathbf{0} >= C_{\mathbf{u}}(\tau) \in \{-v, 1, v+2\}. \tag{33}$$

In the following, we assume that $\mathbf{w} \neq \mathbf{0}$. The difference condition on $\mathbf{e}$ for $\mathbf{u}$ implies that there are at most two zero columns in X, and for the rest of nonzero columns, each has the Hamming weight 2^{m-1}. Note that $\mathbf{w}$ has period v. Therefore, we have the following two facts.

1. There are at most two zero vectors in $\{X_k | 0 \leq k < v\}$ and $H(X_k) = 2^{m-1}$ if $X_k \neq \mathbf{0}$.
2. $W_k \in \{\mathbf{0}, \mathbf{1}\}$. The number of zero vectors in $\{W_k | 0 \leq k < v\}$ is equal to the Hamming weight of $\mathbf{w}$.

Thus $\mathbf{y} = \mathbf{x} + \mathbf{w}$ is a one-shot sequence. Applying Corollary 1 in Section 3 with $R = 2$, $v = t = 2^m - 1$, and $H(\mathbf{w}) \in \{2^{m-1}, 2^{m-1} \pm 2^{\lfloor m/2 \rfloor + 1}\}$, we obtain that $C_{\mathbf{s},\mathbf{t}}(\tau) = I(\mathbf{y}) = \pm j 2^m - (2^m - 1) + 2H(\mathbf{w}), j = 0, 1, 2$. We list possible correlation values in Table 3. (Note. The entries in the second column in Table 3 are from Table 2 in Section 3 where $\mathbf{b}$ is replaced by $\mathbf{w}$.)

Together with (33), correlation values of any pair of sequences, each from IS_i and IS_j, belongs to the following set

$$\{c2^m + e2^{\lfloor m/2 \rfloor + 1} + 1 \,|\, c \in \{0, \pm 1, \pm 2\}, e \in \{0, \pm 1\}\}.$$

Therefore, the maximum magnitude of the correlation values is $\Delta = 2^{m+1} + 2^{\lfloor m/2 \rfloor + 1} + 1$.

□

Table 3. Values of Correlation of IS_i and IS_j

		$C_{\mathbf{s},\mathbf{t}}(\tau)$	
j	$j2^m-(2^m-1)+2H(\mathbf{w})$	$H(\mathbf{w})=2^{m-1}$	$H(\mathbf{w})=2^{m-1}\pm 2^{\lfloor m/2\rfloor}$
0	$-(2^m-1)+2H(\mathbf{w})$	1	$\pm 2^{\lfloor m/2\rfloor+1}+1$
1	$1+2H(\mathbf{w})$	2^m+1	$2^m\pm 2^{\lfloor m/2\rfloor+1}+1$
-1	$-2^{m+1}+1+2H(\mathbf{w})$	-2^m+1	$-2^m\pm 2^{\lfloor m/2\rfloor+1}+1$
2	$2^m+1+2H(\mathbf{w})$	$2^{m+1}+1$	$2^{m+1}\pm 2^{\lfloor m/2\rfloor+1}+1$
-2	$-2^{m+1}-2^m+1+2H(\mathbf{w})$	$-2^{m+1}+1$	$-2^{m+1}\pm 2^{\lfloor m/2\rfloor+1}+1$

Applying Property 1 in Section 1, Theorems 5 and 6, we have the following new interleaved signal set where the slightly increased maximum correlation is traded for a larger size.

Corollary 5. *Let* $IS^+=\cup_{k=0}^{2^m}IS_k$. *Then*

$$IS^+=\{\mathbf{u}+eL^i\mathbf{b}+kL^j\mathbf{c}\,|\,0\le i,j<2^m-1, 0\le e,k\le 1\}$$

and IS^+ *is a* $((2^m-1)^2,2^{2m},\Delta)$ *signal set where* $\Delta=2^{m+1}+2^{\lfloor m/2\rfloor+1}+1$.

For each k, IS_k is an interleaved signal set with parameters $((2^m-1)^2,2^m-1,2^{m+1}+1)$. For the enlarged signal set IS^+, it has parameters $((2^m-1)^2,2^{2m},\Delta)$ where $\Delta=2^{m+1}+2^{\lfloor m/2\rfloor+1}+1$. The size is increased exponentially from 2^m-1 (could be 2^m by adding $\mathbf{u}$ back) to 2^{2m}, which is larger than the period of the sequences in the new signal set, and the maximum correlation increased from $2^{m+1}+1$ to $2^{m+1}+2^{\lfloor m/2\rfloor+1}+1$. Next we look at how much significant that the amount of correlation sacrified is in this approach. The term $2^{\lfloor m/2\rfloor+1}$ is very small compared it with 2^{m+1}, since it is in the order of the square root of 2^{m+1}. Thus when m is large, this degradation of the maximum correlation is insignificant. To support this claim, in Table 4, we list the maximum correlation values of Δ_{IS_k} and Δ_{IS^+} for the signal sets IS_k and IS^+ respectively, expressed in decibels: $10\log_{10}\Delta$, $\Delta\in\{\Delta_{IS_k},\Delta_{IS^+}\}$, for $3\le m\le 22$. From the listed data, the difference between these two maximum correlations in decibel starts at third decimal digit when $m\ge 21$ for m odd and $m\ge 20$ for m even. In other words, they differ maximally only .009 dB when $m\ge 21$ for m odd and $m\ge 20$ for m even.

6 Construction of A Family of Bent Function Signal Sets with Minimum Intraference

We keep the notations of $n=2m$ and $q=2^m$. To give a construction for a family of bent function signal sets and show its maximum correlation is equal to the maximum correlation of an individual bent function signal set, we need some preparations.

Table 4. Comparison of Maximum Correlation in Decibels

m	Δ_{IS_k}, Maximum Correlation of IS_k $10\log_{10}\Delta_{IS_k}$	Δ_{IS^+}, Maximum Correlation of IS^+ $10\log_{10}\Delta_{IS^+}$
odd		
3	12.30448921	13.22219295
5	18.12913357	18.63322860
7	24.09933123	24.36162647
9	30.10723865	30.24074987
11	36.12465964	36.19197716
13	42.14446446	42.17825981
15	48.16486557	48.18179690
17	54.18541579	54.19388980
19	60.20600327	60.21024236
21	66.22660008	66.22872014
even		
4	9.542425094	10.41392685
6	15.18513940	15.68201724
8	21.10589710	21.36720567
10	27.10117365	27.23455672
12	33.11541958	33.18272080
14	39.13442955	39.16822285
16	45.15463188	45.17156295
18	51.17513240	51.18360637
20	57.19570746	57.19994654
22	63.21630116	63.21842122

A bent function $f(x)$ is a function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ whose Hadamard transform has constant magnitude, i.e.,

$$\widehat{f}(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+f(x)} = \pm q, \ \forall \lambda \in \mathbb{F}_{2^n}. \tag{34}$$

Parseval Formula: The Parseval formula states that the correlation of two sequences is equal to the correlation of their Hadamard transforms. In other words, let $g(x)$ be another function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. Then we have

$$C_{\mathbf{a},\mathbf{b}}(\tau) + 1 = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(\delta x)}(-1)^{g(x)} = \frac{1}{2^n} \sum_{y \in \mathbb{F}_{2^n}} \widehat{f}(\delta y)\widehat{g}(y)$$

where $\mathbf{a} = \{a_i\}$ with $a_i = f(\alpha^i)$, $\mathbf{b} = \{b_i\}$ with $b_i = g(\alpha^i)$, and $\delta = \alpha^\tau$.

Composition of Bent Functions: Let $g(x)$ be a function from $\mathbb{F}_q$ to $\mathbb{F}_2$ and $f(x) = g \circ Tr_m^N(x) = g(Tr_m^N(x))$ where $\circ$ is a composition operator, $q = 2^m$ and $N = lm$.

Proposition 1. *With the above notation,*

$$\widehat{f}(\lambda) = \begin{cases} 0, & \lambda \notin \mathbb{F}_q \\ 2^{N-m}\widehat{g}(\lambda), & \lambda \in \mathbb{F}_q. \end{cases}$$

In particular, if $g(x)$ is a bent function from $\mathbb{F}_q$ to $\mathbb{F}_2$, then

$$\widehat{f}(\lambda) = \begin{cases} 0, & \lambda \notin \mathbb{F}_q \\ \pm 2^{N-m/2}, & \lambda \in \mathbb{F}_q. \end{cases}$$

The validity of the proposition can be derived from the results on geometrical sequences discussed by Klapper, Chan and Goresky in [12]. However, for completeness, we provide a proof whose technique will be used in the proof for correlation of the family of bent function signal sets constructed later. To do so, we need the following result about m-sequences generated by $Tr_m^N(x)$ (the proof can be found in [5] or the original paper [24]).

Fact 1 *For $\lambda \in \mathbb{F}_{q^l}$, let*

$$T = \{(Tr_m^N(\lambda x), Tr_m^N(x)) \,|\, x \in \mathbb{F}_{q^l}\}.$$

If $\lambda \notin \mathbb{F}_q$, then each pair $(\theta, \mu) \in \mathbb{F}_q \times \mathbb{F}_q$ occurs q^{l-2} times in T. If $\lambda \in \mathbb{F}_q$, then $(\lambda\mu, \mu)$ occurs q^{l-1} times in T for each $\mu \in \mathbb{F}_q$.

This is referred to as the *2-tuple balance property* of m-sequences over $\mathbb{F}_q$ in [5].

Proof of Proposition 1.

Case 1. $\lambda \notin \mathbb{F}_q$. Applying Fact 1 to the Hadamard transform of $f(x)$, we have

$$\widehat{f}(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+f(x)}$$

$$= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr_1^m(x)\circ Tr_m^N(\lambda x)+g(x)\circ Tr_m^N(x)}$$
$$= q^{l-2} \sum_{\theta,\mu\in\mathbb{F}_{2^m}} (-1)^{Tr_1^m(\theta)+g(\mu)}$$
$$= q^{l-2} \sum_{\theta\in\mathbb{F}_{2^m}} (-1)^{Tr_1^m(\theta)} \sum_{\mu\in\mathbb{F}_{2^m}} (-1)^{g(\mu)} = 0$$

The last identity is equal to zero, because $\sum_{\theta\in\mathbb{F}_{2^m}} (-1)^{Tr_1^m(\theta)} = 0$.

Case 2. $\lambda \in \mathbb{F}_q$. Again, applying Fact 1 to the Hadamard transform of $f(x)$, we have

$$\widehat{f}(\lambda) = q^{l-1} \sum_{\mu\in\mathbb{F}_{2^m}} (-1)^{Tr_1^m(\lambda\mu)+g(\mu)}$$
$$= 2^{N-m}\widehat{g}(\lambda). \tag{35}$$

Thus, the assertion is established. If g is bent, then m is even and $\widehat{g}(\lambda) = \pm 2^{m/2}$. Together with (35), we have

$$\widehat{f}(\lambda) = 2^{N-m}\widehat{g}(\lambda) = \pm 2^{N-m/2}.$$

We are now ready to give a construction for a family of bent function signal sets.

Construction of A Family of Bent Function Signal Sets with Minimum Intraference: Let $f(x) : \mathbb{F}_q \to \mathbb{F}_2$ be bent. We choose $\sigma_0 \in \mathbb{F}_{2^n} \setminus \mathbb{F}_{2^m}$ which is a root of some irreducible polynomial over $\mathbb{F}_q$ of form $x^2 + x + w, w \in \mathbb{F}_q$. Let

$$f_{\lambda,\mu}(x) = f(Tr_m^n(x)) + Tr_1^n((\lambda + \mu\sigma_0)x), \lambda \in \mathbb{F}_q, \mu \in \mathbb{F}_q^*.$$

We define

$$s_{\lambda,\mu,i} = f_{\lambda,\mu}(\alpha^i), i = 0, 1, \cdots.$$

Thus $f_{\lambda,\mu}(x)$ is the trace representation of the sequence $\mathbf{s}_{\lambda,\mu} = \{s_{\lambda,\mu,i}\}_{i\geq 0}$. Let

$$S_\mu = \{\mathbf{s}_{\lambda,\mu} \mid \lambda \in \mathbb{F}_q\}.$$

Then we have a family of signal sets: $\{S_\mu \,|\, \mu \in \mathbb{F}_q^*\}$. For each fixed μ, S_μ is the original bent function signal set with parameters $(2^n - 1, 2^m, 2^m + 1)$ because $\mu\sigma_0 \in \mathbb{F}_{2^n} \setminus \mathbb{F}_q$. Bent function signal sets were constructed by Olsen, Scholtz and Welch in 1982 [18]. The reader is referred to that paper for the details.

Theorem 7. *Any two signal sets S_μ and S_θ are shift distinct when $\mu \neq \theta$.*

Proof. For a pair of sequences $\mathbf{a} \in S_\theta$ and $\mathbf{b} \in S_\mu$, if there is $0 \leq k < 2^n - 1$ such that $\mathbf{b} = L^k\mathbf{a}$, then we have

$$f_{\lambda,\mu}(x) = f_{\eta,\theta}(\delta x), \delta = \alpha^k.$$

which gives that

$$f(Tr_m^n(x)) + Tr_1^n((\lambda + \mu\sigma_0)x) = f(Tr_m^n(\delta x)) + Tr_1^n((\eta + \theta\sigma_0)(\delta x). \tag{36}$$

We can write

$$f(Tr_m^n(x)) = \sum_{i=0}^{2^n-1} h_i x^i, h_i \in \mathbb{F}_{2^n}. \tag{37}$$

Since $f(x)$ is bent, then $f(x)$ is not linear. Therefore, there exists at least one i such that $h_i \neq 0$ and $H(i) > 1$. Together with (36), from one of those i's, applying Lemma 2, we have

$$h_i x^i = h_i \cdot (\delta x)^i \Longrightarrow \delta = 1.$$

Substituting $\delta = 1$ into (36), we obtain

$$Tr_1^n((\lambda + \mu\sigma_0)x) = Tr_1^n((\eta + \theta\sigma_0)(x) \Longrightarrow \lambda + \mu\sigma_0 = \eta + \theta\sigma_0.$$

Since $\{1, \sigma_0\}$ is a basis of $\mathbb{F}_{q^2}$ over $\mathbb{F}_q$, the above identity implies that $\lambda = \eta$ and $\mu = \theta$ which contradicts with $\mu \neq \theta$. Thus S_μ and S_θ are shift distinct as long as $\mu \neq \theta$. □

Theorem 8. *For the family of the bent function signal sets $S_\mu, \mu \in \mathbb{F}_q^*$, any pair of sequences $\mathbf{a}$ and $\mathbf{b}$ where $\mathbf{a} \in S_\theta$ and $\mathbf{b} \in S_\mu$, the magnitude of crosscorrelation between $\mathbf{a}$ and $\mathbf{b}$ is upper bounded by*

$$|C_{\mathbf{a},\mathbf{b}}(\tau)| \leq 2^m + 1, 0 \leq \tau < 2^n - 1$$

where $\tau \neq 0$ if $\mathbf{a} = \mathbf{b}$. Therefore, Δ, the maximum correlation of the family of $2^m - 1$ bent function signal sets $S_\mu, \mu \in \mathbb{F}_q^*$, is given by

$$\Delta = 2^m + 1.$$

Proof. The first part of the proof is similar to that of the original bent function signal set (see [18][13]). If $\tau = 1$, then we have

$$f_{\lambda,\mu}(x) + f_{\eta,\theta}(x) = Tr_1^N(\beta x)$$

where $\beta = \lambda + \mu\sigma_0 + \eta + \theta\sigma_0$. If $\mathbf{a} \neq \mathbf{b}$, then $C_{\mathbf{a},\mathbf{b}}(0) = -1$ if and only if $\beta \neq 0$. Since $\{1, \sigma_0\}$ is basis of $\mathbb{F}_{q^2}$ over $\mathbb{F}_q$, $\beta = 0$ if and only if $\lambda = \eta$ and $\mu = \theta$. Thus $\beta \neq 0$ if and only if $\mathbf{s} \neq \mathbf{t}$. Therefore, we have $C_{\mathbf{a},\mathbf{b}}(0) = -1$ for $\mathbf{a} \neq \mathbf{b}$. In the following, we assume that $\tau \neq 1$.

Using the Parseval formula, we write $\delta = \alpha^\tau$, then

$$|C_{\mathbf{a},\mathbf{b}}(\tau)| = \left| -1 + \sum_{x \in \mathbb{F}_{q^2}} (-1)^{f_{\lambda,\mu}(x)} (-1)^{f_{\eta,\theta}(\delta x)} \right|$$

$$= \left| -1 + \frac{1}{q^2} \sum_{y \in \mathbb{F}_{q^2}} \widehat{f}_{\lambda,\mu}(y) \widehat{f}_{\eta,\theta}(\delta y) \right|$$

$$\leq 1 + \frac{1}{q^2} \max |\widehat{f}_{\lambda,\mu}(y)| \max |\widehat{f}_{\eta,\theta}(\delta y)| \cdot |V|$$

where

$$V = \{y \in \mathbb{F}_{2^n} \,|\, \widehat{f}_{\lambda,\mu}(y) \neq 0 \text{ and } \widehat{f}_{\eta,\theta}(\delta y) \neq 0\}.$$

From Proposition 1, $\max |\widehat{f}_{\lambda,\mu}(y)| = \max |\widehat{f}_{\eta,\theta}(\delta y)| = 2^{n-m/2}$. Substituting to the above equality, we get

$$|C_{\mathbf{a},\mathbf{b}}(\tau)| \leq 1 + \frac{1}{2^n} 2^{n-m/2} 2^{n-m/2} |V| = 1 + 2^m |V|. \tag{38}$$

Thus, we only need to show $|V| \leq 1$. Let

$$T_{\lambda,\mu} = \{y \in \mathbb{F}_{2^n} \,|\, \widehat{f}_{\lambda,\mu}(y) \neq 0\}. \tag{39}$$

From Proposition 1, $\widehat{f}_{\lambda,\mu}(y) \neq 0$ if and only if $y + \lambda + \mu\sigma_0 \in \mathbb{F}_q$. Thus $T_{\lambda,\mu}$ can be rewritten as follows.

$$\begin{aligned} T_{\lambda,\mu} &= \{y \,|\, y + \lambda + \mu\sigma_0 \in \mathbb{F}_q\} \\ &= \{t + \sigma \,|\, t \in \mathbb{F}_q\} \end{aligned}$$

where $t = y + \lambda + \mu\sigma_0$ and $\sigma = \lambda + \mu\sigma_0$. Using this representation, we have

$$\begin{aligned} \delta^{-1} T_{\eta,\theta} &= \{\delta^{-1} y \,|\, \widehat{f}_{\eta,\theta}(y) \neq 0\} \\ &= \{u \,|\, \widehat{f}_{\eta,\theta}(\delta u) \neq 0\}. \end{aligned}$$

Thus V is the intersection of $T_{\lambda,\mu}$ and $\delta^{-1} T_{\eta,\theta}$, i.e.,

$$V = T_{\lambda,\mu} \cap \delta^{-1} T_{\eta,\theta}.$$

For $z \in V$, we have

$$z = t + \sigma = \delta^{-1}(t' + \sigma') \Longrightarrow \delta(t + \sigma) = t' + \sigma'$$

where $t' \in \mathbb{F}_q$ and $\sigma' = \eta + \theta\sigma_0$. Let $\delta = x_0 + x_1\sigma_0, x_0, x_1 \in \mathbb{F}_q$. According to the way we choose σ_0, we have $\sigma_0^2 = \sigma_0 + w, w \in \mathbb{F}_q$. Consequently, we have the following deviation:

$$\begin{aligned} \delta(t + \sigma) &= (x_0 + x_1\sigma_0)(t + \lambda + \mu\sigma_0) \\ &= x_0(t + \lambda) + (x_0\mu + x_1 t + x_1\lambda)\sigma_0 + x_1\mu\sigma_0^2 \\ &= x_0(t + \lambda) + x_1\mu w + [x_0\mu + x_1(t + \lambda + \mu)]\sigma_0 \;\; (\text{using } \sigma_0^2 = \sigma_0 + w) \\ &= t' + \sigma' = t' + \eta + \theta\sigma_0. \end{aligned}$$

From the last two identities, we have

$$x_0\mu + x_1(t + \lambda + \mu) = \theta \tag{40}$$

$$x_0(t + \lambda) + x_1\mu w = t' + \eta. \tag{41}$$

Case 1. $x_1 \neq 0$. In this case, from (40), $t = \frac{\theta + x_0\mu}{x_1} + \lambda + \mu$ and $t' = x_0(t + \lambda) + x_1\mu w + \eta$. Thus V has exactly one element $\Longrightarrow |V| = 1$. Substituting it into (38), we obtain

$$|C_{\mathbf{a},\mathbf{b}}(\tau)| \leq 1 + 2^m. \tag{42}$$

Case 2. $x_1 = 0$. In this case, $\delta = x_0 \in \mathbb{F}_q$. From (40), it follows $x_0 = \mu^{-1}\theta$.

Case 2-a. $\mu = \theta$. We have $x_0 = 1 \Longrightarrow \delta = 1$ which contradicts with $\delta \neq 1$. Thus, $x_1 \neq 0$ if $\mu = \theta$. From Case 1, we have $|C_{\mathbf{a},\mathbf{b}}(\tau)| \leq 1 + 2^m$ for any two distinct sequences in S_μ.

Case 2-b. $\mu \neq \theta$. In this case, we have

$$f_{\lambda,\mu}(x) + f_{\eta,\theta}(\delta x) = f(Tr_m^n(x)) + f(\delta Tr_m^n(x)) + Tr_1^m(Tr_m^n(\beta x))$$

where

$$\beta = \lambda + \mu\sigma_0 + (\eta + \theta\sigma_0)\delta = \lambda + \eta\delta + (\mu + \theta^2\mu^{-1})\sigma_0.$$

The element $\beta \in \mathbb{F}_q$ if and only if

$$\mu + \theta^2\mu^{-1} = 0 \Longrightarrow \mu = \theta.$$

Thus $\beta \notin \mathbb{F}_q$, because $\mu \neq \theta$. Applying Fact 1, it follows that

$$\begin{aligned} C_{\mathbf{a},\mathbf{b}}(\tau) + 1 &= \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(Tr_m^n(x)) + f(\delta Tr_m^n(x)) + Tr_1^m(Tr_m^n(\beta x))} \\ &= \sum_{y,z \in \mathbb{F}_q} (-1)^{f(y) + f(\delta y) + Tr_1^m(z)} \text{ (here } q^{2-2} = 1) \\ &= 0. \end{aligned}$$

This completes the proof.

□

From the proof and the result of Theorem 8, the following corollary is immediate (recall $n = 2m$).

Corollary 6. *Let $BF^+ = \cup_{\mu \in \mathbb{F}_q^*} S_\mu$. Then each sequence in BF^+ has the following trace representation:*

$$f_{\lambda,\mu}(x) = f(Tr_m^n(x)) + Tr_1^n((\lambda + \mu\sigma_0)x), \quad \lambda \in \mathbb{F}_q, \mu \in \mathbb{F}_q^*$$

and BF^+ is a $(2^n - 1, 2^n - 2^m, 2^m + 1)$ signal set.

The family of the bent function signal sets $S_\mu, \mu \in \mathbb{F}_q$ has the minimum maximum intraference, since it retains the same maximum correlation as an individual bent function signal set. For the union set BF^+, it contains the original bent function signal set as a subset, it has $2^n - 2^m$ shift-distinct sequences, and the crosscorrelation between any two sequences in the set or the out-of-phase autocorrelation of any sequence in the set is upper bounded by $1 + 2^m$. Thus, *BF^+ is the best among all known constructions of binary signal sets in terms of the sizes and maximum correlations.*

7 Conclusion and Discussion

We give a mathematical formalization for interference resulted from signals of the other users in CDMA networking systems in the detection process. This type of interference is referred to as intraference among those signal sets. We introduce a new concept of the kth-order shift distinct for a family of signal sets. A family of M signal sets is said to be kth-order shift distinct if any sequence in one of the signal sets is not a linear combination of shifts of sequences from any $k-1$ signal sets choosing from the rest of the signal sets. This case is more realistic for CDMA practice, but it is hard to satisfy. Design of families of kth-order ($k > 2$) shift distinct or pairwise shift distinct signal sets with low correlation constitutes a challenge problem, since it is equivalent to design a signal set with larger size and low correlation which can be decomposed into smaller signal sets. In other words, if we can construct a family of pairwise shift-distinct signal sets with low correlation, then we have a new signal set, which is a union of those signal sets with low correlation and larger size, and vise versa.

We provide a general formula for 0-1 distributions of one-shot sequences or equivalently, the cross-correlation function between two component sequences of a one-shot sequence. This result does not only have important applications in determining correlations of families of Kasami (or generalized Kasami) signal sets and interleaved signal sets, but also in other objectives of sequence design.

In the following, we summarize the results on three families of signal sets and their corresponding union sets that we presented in this paper in Table 5. For comparison, we also list the parameters for an individual signal set in each family of the signal sets.

Among three families, listed in Table 5, only the Kasami signal sets which consists of m Kasami signal sets with period $2^{2m} - 1$ satisfies the mth-order shift-distinct property. The other two families of the signal sets only satisfy the pairwise shift-distinct property. However, the maximum correlation that we derived for the family of the Kasami signal sets is only for any two "consecutive" signal sets. This means that for detecting a signal from the signal set S_i, if the interference that the detector experienced from other signals in the communication range is from either S_{i+1} or S_{i-1}, then the interference could be low, since the maximum correlation between S_{i-1} and S_i or S_{i+1} and S_i are low. For both families of

Table 5. Properties of Families of Signal Sets and Their Union Signal Sets

Parameters	Family of the Signal Sets			
	Kasami	Generalized Kasami	Interleave	Bent Function
Period	$2^{2m}-1$		$(2^m-1)^2$	$2^{2m}-1$
Number of the Signal Sets in One Family	m		2^m+1	2^m-1
Maximum Correlation Δ	$2^{m+1}+2^m-1$		$2^{m+1}+2^{\lfloor m/2\rfloor+1}+1$	2^m+1
kth-order Shift distinct, $k>2$	Yes, $k=m$	Not clear	Not clear	Not clear
Pairwise Shift distinct	Yes	Possibly	Yes	Yes
Networking Topology: Δ is for	any two "consecutive" signal sets		any two signal sets	any two signal sets
	Individual Signal Set			
Period	$2^{2m}-1$		$(2^m-1)^2$	$2^{2m}-1$
Size	2^m		2^m-1	2^m
Maximum Correlation	2^m+1		$2^{m+1}+1$	2^m+1
	Derived New Signal Sets: Union Sets			
	Kasami K^+		Interleave IS^+	Bent Function BF^+
Period	$2^{2m}-1$		$(2^m-1)^2$	$2^{2m}-1$
Size	2^{3m}		2^{2m}	$2^m(2^m-1)$
Maximum Correlation	$2^{m+1}+2^m-1$		$2^{m+1}+2^{\lfloor m/2\rfloor+1}+1$	2^m+1

interleaved signal sets and bent function signal sets, they satisfy the pairwise shift-distinct property (it is not clear whether they satisfy mth-order shift-distinct property for $m > 2$). However, they both have very large sizes for each family of these signal sets: there are $2^m + 1$ pairwise shift-distinct interleaved signal sets with period $(2^m - 1)^2$ and $2^m - 1$ pairwise shift-distinct bent function signal sets with period $2^{2m} - 1$ in each family, respectively. The most important feature of the last two families of signal sets, compared with the family of the Kasami signal sets, is that the maximum correlation of each of the both families is designed to be low for any two signal sets in the family. Hence, in the detection process, the intraference resulted from any other signals in the communication range is low for both cases. The other remarkable findings here are that (1) the maximum correlation of the family of the interleaved signal sets is comparable with an individual interleaved signal set, and (2) the maximum correlation of the family of bent function signal sets is minimized, since it is equal to an individual bent function signal set in the family.

From those three families of the signal sets, we obtain three new signal sets. Among them, the enlarged bent function signal set BF^+ has minimized maximum correlation and contains the original bent function signal set as a subset. In other words, BF^+ maintains the same maximum correlation as the original bent function signal set and has size which is almost the square of that of the original bent function signal set. On the other hand, the set BF^+ has the same maximum correlation and period as those of the Kasami signal set, but the ratio of the number of sequences in the set to the period of the sequences is slightly smaller than that of the Gold-pair signal set, because the former has ratio $\frac{2^m(2^m-1)}{2^{2m}-1}$ and the latter has ratio $\frac{2^{2m+1}+1}{2^{2m+1}-1}$. Both will be converged to 1 for a large m. However, the maximum correlation of the Gold-pair signal set is $1 + 2^{m+1}$ which is worse than that of BF^+ which is $1 + 2^m$. Thus, BF^+ is the *best known* binary signal set in terms of the number of signals in the set and the maximum correlation. The maximum correlation of the new interleaved signal set IS^+ is comparable with the original interleaved signal set, whose correlation is maximally degraded by .009 in decibel (dB) when $m \geq 21$ for m odd and $m \geq 20$ for m even. The size of enlarged interleaved signal set is 2^{2m} which is larger than the period of the sequences, and the ratio of the size to the period is greater than that of the Gold-pair signal set. The enlarged Kasami signal set K^+ has the largest size among these three new sets, but the correlation is degraded more than 3 in dB, compared with the original Kasami signal set.

Note that for the family of the interleaved signal sets, the family size could be further enlarged while the maximum correlation is not degraded significantly using other signal sets with low correlation and large size, for example, one may use the enlarged Kasami signal set K^+ as an option for selecting sequence $\mathbf{b}$ in the interleaved construction.

The linear spans of sequences in the enlarged bent function signal set are the same as those of the original bent function signal set (see Kumar's work in 1984 [13] on linear span of the bent function sequences), the linear spans of sequences in the enlarged interleaved signal set are greater than that of the original interleaved signal set by m, and the linear spans of the enlarged Kasami signal set are greater than the Kasami signal set by $2m$. The implementations for the three enlarged signal sets are roughly the same as their individual signal sets.

The last result presented in this paper is for the family of generalized Kasami signal sets. The only result that we obtained here is the maximum correlation, which is the same as that of the family of Kasami signal sets. For the problems on the kth-order shift-distinct property for $2 < k \leq m$ and how to construct an enlarged generalized Kasami signal set, they remain unsolved.

References

1. T.W. Cusick and H. Dobbertin, Some new three-valued crosscorrelation functions for binary m-sequences, *IEEE Trans. on Inform. Theory*, Vol. 42, No. 4, July 1996, pp. 1238-1240.
2. H. Dobbertin, G. Leander, A. Canteaut, C. Carlet, P. Felke, and P. Gaborit, Construction of bent functions via Niho power functions, appear soon at *Journal of Combinatorial Theory, Series A*.
3. R. Gold, Maximal recursive sequences with 3-valued recursive cross-correlation functions, *IEEE Trans. on Inform. Theory*, Vol. 14, January 1968, pp. 154-156.
4. S.W. Golomb, *Shift Register Sequences,* Holden-Day, Inc., San Francisco, 1967, revised edition, Aegean Park Press, Laguna Hills, CA, (1982).
5. S.W. Golomb and G. Gong, *Signal Design with Good Correlation: for Wireless Communications, Cryptography and Radar Applications*, Cambridge University Press, 2005.
6. G. Gong, Correlation among signal sets, presented at the Special Session on *Coding Theory and Cryptography*, organized by Horacio Tapia-Recillas and Neal Koblitz, at *VI Joint Meeting of American Mathematical Society and Sociedad Matematica Mexicana (AMS-SMM)*, May 12-15, 2004, Huston.
7. G. Gong, Theory and applications of q-ary interleaved sequences, *IEEE Trans. on Inform. Theory*, Vol. 41, No. 2, March 1995, pp. 400-411.
8. G. Gong, New Designs for signal sets with low cross-correlation, balance property and large linear span: $GF(p)$ Case, *IEEE Trans. on Inform. Theory*, Vol. 48, No.11, November 2002, pp. 2847-2867.
9. G. Gong, Array correlation and sequences with equal magnitude correlation, *Mathematical Properties of Sequences and Other Combinatorial Structures*, J. -S. No, H. -Y. Song, T. Helleseth, and V. Kumar (Ed), Kluwer Academic Publishers, New York, 2003, pp.33-43.
10. T. Kasami, The weight enumerators for several classes of subcodes of the 2nd-order Reed-Muller codes, *Information and Control*, Vol. 18, pp. 369-394, 1971.
11. 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, pp. 1765-1853.
12. A. Klapper, A.H. Chan, and M. Goresky, Cross-correlations of linearly and quadratically related geometric sequences and GMW sequences, *Discrete Appl. Math.* **46**, No. 1, (1993), pp. 1-20.
13. P.V. Kumar, On bent sequences and generalized bent functions, Ph. D. Thesis, University of Southern California, Los Angeles, 1983.
14. R. Lidl and H. Niederreiter, *Finite Fields*, Encyclopedia of Mathematics and its Applications, Volume 20, Addison-Wesley, (1983). (Revised version, Cambridge University Press, 1997.)
15. R.J. McEliece, *Finite Fields for Computer Scientists and Engineers*, The Knuwer International Series in Engineering and Compute Science, Vol. 23, Kluwer Academic Publishers, Boston, (1986).
16. J.S. No, K. Yang, H.G. Chung and H.Y. Song, New constructions for families of binary sequences with optimal correlation properties, *IEEE Trans. Inform. Theory*, Vol. 43, No. 5, September 1997, pp. 1596-1602.

17. J.S. No and P.V. Kumar, A new family of binary pseudo-random sequences having optimal periodic correlation properties and larger linear span, *IEEE Trans. Inform. Theory*, vol 35, No. 2, March 1989, pp. 371-379.
18. J.D. Olsen, R.A. Scholtz and L.R. Welch, Bent-function sequences, *IEEE Trans. Inform. Theory*, Vol. 28, No. 6, November 1982, pp. 858-864.
19. K.G. Paterson, Binary sequence sets with favorable correlations from difference sets and MDS codes, *IEEE Trans. Inform. Theory*, Vol. 44, No. 1, January 1998, pp. 172-180.
20. M.B. Pursley, *Introduction to Digital Communications*, Pearson Prentice Hall, 2005.
21. M.K. Simon, J.K. Omura, R.A. Scholtz, and B.K. Levitt, *Spread Sprectrum Communications Handbook*, McGraw-Hill Companies, Inc., 2002.
22. A.J. Viterbi, *CDMA − Principles of Spread Spectrum Communication*, Reading, MASS., Addison-Wesley, 1995.
23. A. M. Youssef and G. Gong, Hyper-Bent Functions, *Advances in Cryptology, EuroCrypto 2001, Lecture Notes in Computer Science*, Brigit Pfitzmann (Ed). Berlin, Springer-Verlag, 2001, vol. 2045, pp.406-419.
24. N. Zierler, Linear recurring sequences, *J. Soc. Indust. Appl. Math.* **7** , (1959), pp. 31-48.