World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:2, No:12, 2008

On Pseudo-Random and Orthogonal Binary Spreading Sequences

Abhijit Mitra

Abstract**—**Different pseudo-random or pseudo-noise (PN) as well as orthogonal sequences that can be used as spreading codes for code division multiple access (CDMA) cellular networks or can be used for encrypting speech signals to reduce the residual intelligence are investigated. We briefly review the theoretical background for direct sequence CDMA systems and describe the main characteristics of the maximal length, Gold, Barker, and Kasami sequences. We also discuss about variable- and fixed-length orthogonal codes like Walsh-Hadamard codes. The equivalence of PN and orthogonal codes are also derived. Finally, a new PN sequence is proposed which is shown to have certain better properties than the existing codes.

Keywords**—**Code division multiple access, pseudo-noise codes, maximal length, Gold, Barker, Kasami, Walsh-Hadamard, autocorrelation, crosscorrelation, figure of merit.

I. Introduction

IN a spread spectrum CDMA scheme, the transmitted bandlimited signal is spread over a wide frequency band, much wider than the minimum bandwidth required to transmit the information being sent [1]. It employs a waveform that for all purposes appears random to anyone but the intended receiver of the transmitter waveform. Actually, for ease of both generation and synchronization by the receiver, the waveform is taken as a *random-like*, meaning that it can be generated by mathematically precise rules, but statistically it satisfies the requirements of a truly random sequence in the limiting sense. These pseudo-random or pseudo-noise (PN) properties include, among other properties, balance, run and autocorrelation properties. In spread spectrum CDMA all users use the same bandwidth, but each transmitter is assigned a distinct code. For using same bandwidth, there lies a possible chance of high interference among the users when all are close. The spreading codes, therefore, play a vital role in designing such a system where each one should have high autocorrelation and very less, ideally zero, crosscorrelation values. A direct sequence (DS) CDMA modulation technique using a PN code is shown in Fig. 1. Multiplying the input data by a PN sequence, the bit rate of which is much higher than the data bit rate, is called *spreading*. This increases the data rate while adding redundancy to the system. The ratio of PN sequence bit rate to data bit rate is called the *spreading factor* (SF). The resulting waveform is wideband, noise-like, and balanced in phase. In case there are two different I and Q branches, each channel can be spread separately (which is usually done in wideband CDMA). When the signal is received, the spreading is removed from the received code by multiplying the same PN code exactly synchronized with the received signal. However, if the CDMA codes are designed to have very low crosscorrelation, then there is no despreading generated by other user's interference signals.

The notion of PN codes can be used for encryption purpose as well. While transmitting information (speech, image or other data) through insecure channels, there might be unwanted disclosure as well as unauthorized modification of data if that is not properly secured. Therefore, certain mechanisms are needed to protect the information within insecure channel. One way to provide such protection is to convert the intelligible data into unintelligible form prior to transmission and such a process of conversion with a key is called encryption [2]. At the receiver side, the encrypted message is converted back to the original intelligible form by the reverse process of the encryption called decryption.

For these two important application areas, PN sequences have received a lot of attention lately. We take an analytical approach in this article to show the generation, code length, autocorrelation and crosscorrelation properties of some important PN [3]-[15] and orthogonal sequences [16]-[19]. The main among them are maximal length, Gold, Barker, Kasami, and Walsh-Hadamard sequences. We also provide the mean square correlation measurements for evaluating the performance of different PN sequences. We next propose a new class of PN sequence and show that it has better correlation properties than the existing codes that makes the code a potential alternate choice as spreading codes in CDMA systems.

This paper is organized as follows. In Sections 2 and 3, we deal with different kind of PN and orthogonal sequences. Section 4 shows the equivalence of PN and orthogonal codes. Section 5 describes the main techniques to measure the correlation properties of the PN sequences practically. We introduce the new class of PN sequence with one of its important properties in Section 6. The paper is concluded by summarizing the main concepts discussed herein in Section 7.

II. Fundamentals of Different PN sequences

PN sequences are sequence of 1's and 0's where the numbers look like statistically independent and uniformly distributed. As said earlier, they are arranged *random-like*, meaning that it can be generated by mathematically precise rules, but statistically it satisfies the requirements of a truly random sequence in the limiting sense. The PN sequences have the following noise like properties [5]:

Manuscript received July 07, 2007.

The author is with the Department of Electronics and Communication Engineering, Indian Institute of Technology (IIT) Guwahati, India. E-Mail: a.mitra@iitg.ernet.in.

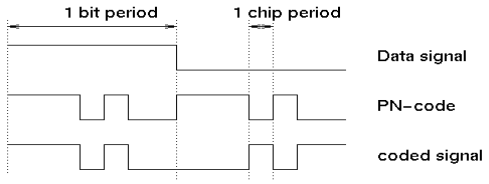


Fig. 1. A DS-CDMA modulation technique using a PN code.

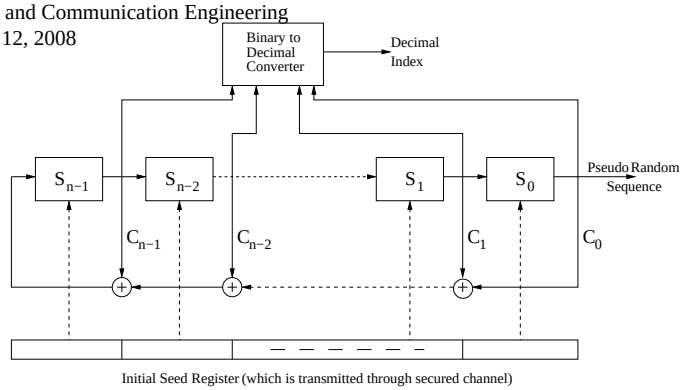


Fig. 2. The block diagram of an m-sequence generator.

(i) Balance: Any PN sequence of length $N = 2^n - 1$ contains 2^{n-1} ones and $2^{n-1} - 1$ zeros,
(ii) Run: In any PN sequence, $1/2$ of the runs have length 1, $1/4$ have length 2, $1/8$ have length 3, $1/16$ have length 4, and so on, as long as these fractions give integral numbers of runs (in each case, the number of runs of 0's is equal to the number of run's of 1's),
(iii) Autocorrelation: The autocorrelation function $r(i)$ of any PN sequence of length N is given by

$$r(i) = \begin{cases} 1 & \text{for } i = 0 \\ -\frac{1}{N} & \text{for } 1 \leq |i| \leq N-1. \end{cases}$$

These three properties make PN sequences efficient for speech encryption. However, particularly due to the third property, adjacent bits correlation becomes considerably less, thereby making the PN sequences more effective to be used in systems like CDMA. Therefore, useful PN sequences must have very good auto-correlation and cross-correlation properties as well as maintaining some randomness properties. The Welch bound places a lower limit on the maximum level of the correlation function (auto-correlation of sidelobes and cross-correlation levels). The Welch bound for a set of K sequences with each sequence of length N ($N \geq K$) is defined as

$$\phi_{max} \geq \sqrt{\frac{N-K}{NK-K}} \tag{1}$$

and such a bound is no longer achievable when $N > K(K+1)/2$ for real cases. Note that, in the sequel, sometimes we shall represent binary sequences using zeros and ones and in other cases +1's and -1's. The appropriate mapping is that the zeros are mapped to +1's and ones are mapped to -1's. Below, we describe certain main PN sequences.

A. Maximal Length Sequences

The Maximal length sequence (m-sequence) generator is usually constructed with linear feedback shift registers (LFSR) [8],[9]. The m-sequences are, by definition, the largest codes that can be generated by a given shift register of given length with feedback. The feedback function, also called as characteristic polynomial, determines the length and type of the sequence generated. Say, $p(x) = x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_0$. In a finite field of or Galois field ($GF(q)$) over a prime number q, $p(x)$ is irreducible if it cannot be factored into a product of lower degree polynomials. It becomes primitive if the smallest integer l for which $p(x)$ divides $c(x) = x^l - 1$ is $l = q^n - 1$. Fig. 2 shows the general structure of a m-sequence generator. It contains n shift registers and is initiated with a starting seed, which is usually transmitted to the intended users through a secure channel. The outputs of the shift registers are multiplied with the coefficients $(C_{n-1}, C_{n-2}, \ldots, C_1, C_0)$ of a primitive polynomial with respect to modulo-2 operation. The resultant output obtained by the modulo-2 operation is then fed back to the first shift register and is called an m-sequence. The length of the generated m-sequence is $N = 2^n - 1$, which depends on the length of the LFSR used for the generation of the m-sequence. For binary fields ($GF(2^n)$), a set of primitive polynomials up to $n = 30$ is shown in Table 1, which is quite sufficient for most of the purposes. Primitive polynomials of much higher degree can be found in [18].

An m-sequence possesses good randomness properties including a two valued auto-correlation function: level one which occurs for every N elements, when all the patterns match, and has a correlation of 1, and level two has a correlation of $-\frac{1}{N}$. In other words,

$$\mathbf{r}(\tau) = \begin{cases} 1 & \text{for } \tau = 0, N, 2N, \ldots \\ -\frac{1}{N} & \text{otherwise.} \end{cases} \tag{2}$$

B. Gold Sequences

Gold sequences are generated by the modulo-2 operation of two different m-sequences of same length. Fig. 3 shows the block diagram of a Gold sequence generator. The two m-sequences are able to generate a family of many non maximal product codes, but a preferred maximal sequences can only produce Gold codes [4]. Finding preferred pair of m-sequences is necessary in defining set of Gold sequences. Since both m-sequences have equal length N, the generated gold sequence is of length N as well. From a pair of preferred sequences, the Gold sequences are generated by the modulo-2 sum of the first with shifted versions of the second or vice-versa. For a period of $N = 2^n - 1$, there are N possible circular shifts. Thus, one can get N sequences with two preferred m-sequences, and these are called Gold sequences [5].

Consider an m-sequence represented by a binary vector u of length N, and a second sequence v obtained by sampling every q^{th} symbol of u. In other words, $v = u[q]$, where q

TABLE I
A SET OF PRIMITIVE POLYNOMIALS UP TO DEGREE 30 FOR FOR GENERATING m-SEQUENCES

degree (n)	$p(x)$
1	$x+1$
2	x^2+x+1
3	x^3+x+1
4	x^4+x+1
5	x^5+x^2+1
6	x^6+x+1
7	x^7+x+1
8	$x^8+x^6+x^5+x+1$
9	x^9+x^4+1
10	$x^{10}+x^3+1$
11	$x^{11}+x^2+1$
12	$x^{12}+x^7+x^4+x^3+1$
13	$x^{13}+x^4+x^3+x+1$
14	$x^{14}+x^{12}+x^{11}+x+1$
15	$x^{15}+x+1$
16	$x^{16}+x^5+x^3+x^2+1$
17	$x^{17}+x^3+1$
18	$x^{18}+x^7+1$
19	$x^{19}+x^6+x^5+x+1$
20	$x^{20}+x^3+1$
21	$x^{21}+x^2+1$
22	$x^{22}+x+1$
23	$x^{23}+x^5+1$
24	$x^{24}+x^4+x^3+x+1$
25	$x^{25}+x^3+1$
26	$x^{26}+x^8+x^7+x+1$
27	$x^{27}+x^8+x^7+x+1$
28	$x^{28}+x^3+1$
29	$x^{29}+x^2+1$
30	$x^{30}+x^{16}+x^{15}+x+1$

is odd and either $q = 2^k + 1$ or $q = 2^{2k} - 2^k + 1$. Two m-sequences u and v are called the preferred pair if

$$n \neq 0 \pmod 4 \tag{3}$$

i.e., n is odd or $n = 2 \pmod 4$. The relation between n and k in such Gold sequences follows the below stated property.

$$\gcd(n,k) = \begin{cases} 1 & \text{for } n \text{ odd} \\ 2 & \text{for } n = 2 \pmod 4. \end{cases} \tag{4}$$

The set of Gold sequences generated with the two preferred pair of m-sequences u and v is given as:

$$\mathbf{G}(u,v) = \{u, v, u \oplus v, u \oplus T \cdot v, u \oplus T^2 \cdot v, \ldots, u \oplus T^{N-1} \cdot v\} \tag{5}$$

where T is the cyclic shift operator and $\oplus$ is the XOR operation.

The auto-correlation properties of Gold sequences are not as good as that of m-sequences. Apart from the two original sequences, the other are not m-sequences; hence the auto-correlation is not two valued. The Gold sequences set have three valued auto-correlation spectrum, but these sequences provides more security compared to m-sequences. Gold sequences auto-correlation function $\mathbf{r}_{xx}(\tau)$ and the cross-correlation function $\mathbf{r}_{xy}(\tau)$ can be defined as

$$\mathbf{r}_{xx}(\tau) = \begin{cases} 1 & \text{for } \tau = 0 \\ \left\{-\frac{t(n)}{N}, -\frac{1}{N}, \frac{t(n)-2}{N}\right\} & \text{for } \tau \neq 0 \end{cases} \tag{6}$$

$$\mathbf{r}_{xy}(\tau) = \left\{-\frac{t(n)}{N}, -\frac{1}{N}, \frac{t(n)-2}{N}\right\} \tag{7}$$

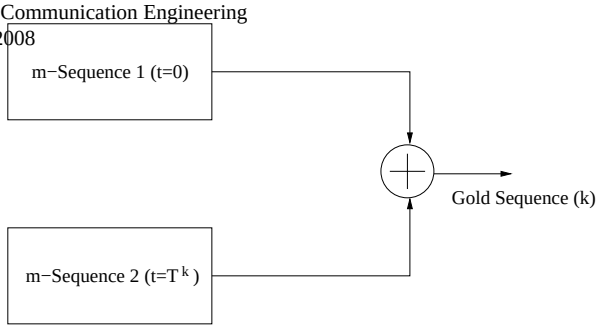


Fig. 3. A typical Gold sequence generator.

where

$$t(n) = \begin{cases} 1 + 2^{\frac{n+1}{2}} & \text{for } n \text{ odd} \\ 1 + 2^{\frac{n+2}{2}} & \text{for } n \text{ even.} \end{cases}$$

The peak correlation value for Gold sequences is $t(n)/N$ and from the above equations, Gold sequences with odd value of LFSR length have less correlation values compared to that of even value of n, since the value of $t(n)$ for even value of n is less.

C. Gold-Like Sequences

There exist a class of sequences which have parameters similar to those of Gold sequences. Let u be an m-sequence of length $N = 2^n - 1$ generated by a primitive polynomial of degree n and let q be an integer such that $\gcd(q, N) = 3$ and let $v^{(k)}$, $k = 0, 1, 2$, denote the sequences obtained by decimating $T^k u$ by q. The sequences $v^{(k)}$ are periodic with period $N' = N/3$. This new class of sequences, defined as $\mathbf{G}_q(u,v) = \{u, u \oplus v^{(0)}, u \oplus Tv^{(0)}, u \oplus T^1 v^{(0)}, \cdots, u \oplus T^{N'-1} v^{(0)}, u \oplus v^{(1)}, u \oplus Tv^{(1)}, u \oplus T^1 v^{(1)}, \cdots, u \oplus T^{N'-1} v^{(1)}, u \oplus v^{(2)}, u \oplus Tv^{(2)}, u \oplus T^1 v^{(2)}, \cdots, u \oplus T^{N'-1} v^{(2)}\}$, are called Gold-like sequences [10].

The Gold-like sequence set contains 2^n sequences and period of the each sequence is $2^n - 1$. The peak correlation value for the set $\mathbf{G}_q(u,v)$ is $\phi_{max} = t(n)$, which is similar to that of the Gold sequences.

D. Barker Sequences

Barker sequences are short length codes that offer good correlation properties. A Barker code is a sequence, $\{\mathbf{c}_i\}$, of 1's and 0's of some finite length N such that the discrete auto-correlation function $\mathbf{r}(\tau)$, defined as:

$$\mathbf{r}(\tau) = \sum_{i=0}^{N-\tau} \mathbf{c}_i \mathbf{c}_{i+\tau} \tag{8}$$

satisfies $|\mathbf{r}(\tau)| \leq 1$ for $(\tau \neq 0)$ [11]. Barker sequences have many advantages over other PN sequences. These sequences satisfies the balance property and the run property of the PN sequences. These sequences have uniformly low auto-correlation sidelobes ($\leq$1), and the size of these families is also small. Table II gives a list of known Barker sequences. No theoretical limit to the length of Barker sequences has been found. Turyn and Storer [12] have shown that there are

TABLE II
A LIST OF KNOWN BARKER SEQUENCES

Length (bits)	Sequence
2	[1 -1], [1 1]
3	[1 1 -1]
4	[1 -1 1 1], [1 -1 -1 -1]
5	[1 1 1 -1 1]
7	[1 1 1 -1 -1 1 -1]
11	[1 1 1 -1 -1 -1 1 -1 -1 1 -1]
13	[1 1 1 1 1 -1 -1 1 1 -1 1 -1 1]

no other solutions of odd length and very little has been said concerning the existence of Barker codes of even length. In [13] it has been shown that codes of even length must have a length which is a perfect square (except $N = 2$).

E. Barker-Like Sequences

Barker sequences have good correlation properties with the peak correlation value of these sequences being bounded by 1. The number of existing Barker sequences, however, are very less. We can generate more number of sequences by making some relaxation on the peak value of the correlation function. These new sequences are called Barker-like sequences [14] which are defined as follows:

$$\mathbf{r}(\tau) \leq m \qquad 0 < \tau \leq \tau' < N \tag{9}$$

where $\mathbf{r}(\tau)$ is the auto-correlation function for a sequence of length N as defined in (8). The variables τ' and m are design parameters for the Barker like sequences, where τ' is the maximum allowed shift between the sequence and m is the magnitude of the upper bound on the peak correlation function. If $m = 1$ and $\tau' = N$, the generated sequences becomes Barker sequences.

F. Kasami Sequences

Kasami sequences are PN sequences of length $N = 2^n - 1$, where n is the degree of the primitive polynomial used to generate Kasami sequences and these sequences are defined for even values of n [10],[15]. There are two classes of Kasami sequences: (i) small set of Kasami sequences, (ii) large set of Kasami sequences.

1) Small Set of Kasami Sequences: Small set of Kasami sequences are optimal in the sense of matching Welch's lower bound for correlation functions. A small set of Kasami sequences [10] is a set of $2^{n/2}$ binary sequences each of length $N = 2^n - 1$, where n is the degree of the higher order primitive polynomial used to generate the set and it is of even length. The degree of the second order polynomial used to generate small Kasami set is $n/2$. Let u be an m-sequence of length N generated by a primitive polynomial of length n, and let w be the sequence obtained by decimating u by $2^{n/2} + 1$. The small set of Kasami sequences is defined as:

$$\mathbf{k}_s(u) = \left\{u, u \oplus w, u \oplus Tw, \cdots, u \oplus T^{2^{n/2}-2}w\right\}. \tag{10}$$

The sequence w obtained by decimating the m-sequence u by $2^{n/2} + 1$ is also an m-sequence of length $2^{n/2} - 1$, generated by a primitive polynomial of degree $n/2$. Similar to the Gold codes the correlation function of small set of Kasami sequences is also three valued. The correlation functions for the sequences take on from the values $\{-1, -s(n), s(n) - 2\}$, where $s(n) = 2^{n/2} + 1$. The auto-correlation function of small set of Kasami sequences is defined as:

$$\mathbf{r}_{xx}(\tau) = \begin{cases} 1 & \text{for } \tau = 0 \\ \left\{-\frac{s(n)}{N}, -\frac{1}{N}, \frac{s(n)-2}{N}\right\} & \text{for } \tau \neq 0. \end{cases} \tag{11}$$

2) Large Set of Kasami Sequences: Small set of Kasami sequences are optimal sequences, these sequences have better correlation properties compared to Gold sequences. However, the set contains less number of sequences. For the shift register of length n the number of possible sequences for the small Kasami sequence set is only $2^{n/2}$ sequences, whereas a Gold code set contains $2^n + 2$ sequences. The number of sequences can be increased by making some relaxation on the correlation values of the sequences. The new set of sequences is called large set of Kasami sequences [10],[15]. The large set of Kasami sequences contains more number of sequences compared to Gold codes but these sequences have more correlation values compared to Gold codes. The large set of Kasami sequences contains all the sequences in the small set along with the Gold codes.

For $\text{mod}(n, 4) = 2$, the large set of Kasami sequences is defined as follows. Let v be the sequence formed by decimating the sequence u by $t(n)$. The large set of Kasami sequences is then defined as:

$$\mathbf{k}_l(u) = \mathbf{G}(u, v) \bigcup \left[\bigcup_{i=0}^{2^{n/2}-2} \left\{T^i w \oplus G(u, v)\right\} \right]. \tag{12}$$

The correlation functions for the sequences takes on the values $\{-t(n), -s(n), -1, s(n) - 2, t(n) - 2\}$. The maximum value of the correlation function is $t(n)$ which is same as that of Gold codes for even value of n and for odd value of n the Gold codes have less correlation values compared to large set of Kasami sequences. Table III provides a comparative account of the length of the PN sequences, number of possible PN sequences of a class and the peak correlation values for different PN sequences depending on the degree of the primitive polynomial used to generate the sequences.

III. ORTHOGONAL CODES

Two sequences are said to be orthogonal when the inner product between the two sequences is zero. If $\mathbf{c}_i(k\tau)$ and $\mathbf{c}_j(k\tau)$ are the i^{th} and j^{th} orthogonal members of an orthogonal set, respectively, M is the length of the set and τ is the symbol duration, then the orthogonal property states that

$$\sum_{k=0}^{M-1} \mathbf{c}_i(k\tau)\mathbf{c}_j(k\tau) = 0 \qquad i \neq j. \tag{13}$$

TABLE III
A COMPARISON OF FAMILIES OF PN SEQUENCES

Sequence	n	Length	Number of Sequences	ϕ_{max}
Maximal	even/odd	2^n-1	1	1
Gold	odd	2^n-1	2^n-1	$1+2^{\frac{n+1}{2}}$
Gold	even	2^n-1	2^n-1	$1+2^{\frac{n+2}{2}}$
Gold-Like	$\gcd(n,k)=3$	2^n-1	2^n	$1+2^{\frac{n+2}{2}}$
Barker-Like	even/odd	2^n-1	$f(n,m)$	m
Small Kasami	even	2^n-1	$2^{\frac{n}{2}}$	$1+2^{\frac{n+2}{2}}$
Large Kasami	even and $\mod(n,4)=2$	2^n-1	$[2^{\frac{n}{2}}(2^n+1)-1]$	$1+2^{\frac{n+2}{2}}$

There are two kinds of orthogonal codes: fixed- and variable-length, which obey this relation and are discussed below.

A. Fixed Length Orthogonal Codes

1) Walsh Hadamard Codes: Walsh Hadamard (WH) codes are orthogonal codes possessing low auto-correlation properties. The WH sequences [16], [17] of length N are defined with a class of orthogonal matrices $\mathbf{H}_N$, called Hadamard matrices, as follows

$$\mathbf{H}_N\mathbf{H}_N^T = N\mathbf{I}_N \tag{14}$$

where $\mathbf{H}_N^T$ is the transposed Hadamard matrix of order N, and $\mathbf{I}_N$ is the $N\times N$ unity matrix.

Walsh sequences are the rows of a Hadamard matrix $\mathbf{H}_N$, which is a square matrix of order N. Hadamard matrices contain one row of all ones and the remaining rows each have equal numbers of +1's and -1's. Walsh sequences can be constructed for block length $N=2^n$, where n is an integer. The Hadamard matrix of desired length can be generated by the following recursive procedure

$$\mathbf{H}_1 = \begin{bmatrix} 1 \end{bmatrix}_{1\times 1},\ \mathbf{H}_2 = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_1 \\ \mathbf{H}_1 & -\mathbf{H}_1 \end{bmatrix}_{2\times 2}, \cdots$$

$$\mathbf{H}_N = \begin{bmatrix} \mathbf{H}_{N/2} & \mathbf{H}_{N/2} \\ \mathbf{H}_{N/2} & -\mathbf{H}_{N/2} \end{bmatrix}_{N\times N}.$$

From the corresponding matrix, the WH sequences are given by the rows, in Walsh functions the 0's are mapped to 1's and 1's are mapped to -1's. However, the auto-correlation function of WH codes does not have good characteristics, it can have more than one peak and these sequences do not satisfy the run property.

2) Modified Walsh Hadamard Codes: Modified Walsh Hadamard (MWH) codes are generated by multiplying the Hadamard matrix $\mathbf{H}_N$ with a diagonal matrix $\mathbf{D}_N$ of same order [17],[19] as follows

$$\mathbf{W}_N = \mathbf{H}_N\mathbf{D}_N. \tag{15}$$

The modified WH codes are orthogonal codes too, since

$$\mathbf{W}_N\mathbf{W}_N^T = \mathbf{H}_N\mathbf{D}_N(\mathbf{H}_N\mathbf{D}_N)^T = \mathbf{H}_N\mathbf{D}_N\mathbf{D}_N^T\mathbf{H}_N^T \tag{16}$$

where it can be shown that

$$\mathbf{D}_N\mathbf{D}_N^T = k\mathbf{I}_N \qquad k\in\mathcal{R}. \tag{17}$$

Substituting (17) into (16) yields

$$\mathbf{W}_N\mathbf{W}_N^T = k\mathbf{H}_N\mathbf{H}_N^T = kN\mathbf{I}_N. \tag{18}$$

If $k=1$, then the sequences defined by the matrix $\mathbf{W}_N$ are not only orthogonal, but possess the same normalization as the WH sequences. However, the correlation properties of the sequences defined by $\mathbf{W}_N$ can be significantly different to those of the original WH sequences. These sequences have better correlation properties compared to Walsh sequences. The simple class of orthogonal matrices that can be chosen is diagonal matrices, with their elements ($d_{m,n}$) fulfilling the condition:

$$d_{m,n} = \begin{cases} 0 & \text{for } m\neq n \\ k & \text{for } m=n; \quad m,n=1,2,\ldots,N \end{cases}$$

where $|k|=1$.

B. Variable Length Orthogonal Codes

WH codes and MWH codes are fixed length orthogonal codes with the dot product of any two such sequences being zero. Variable length orthogonal codes $\mathbf{c}_k(i)$, $i=1,2,\ldots,2^k$, are those with different lengths satisfying the orthogonal property [4]. The codes are taken from a orthogonal variable spreading factor (OVSF) code tree as shown in Fig. 4. This code tree generation algorithm is similar to the recursive generation of the Walsh codes by means of the Hadamard matrices. New levels in the code tree are generated by concatenating a root codeword with a replica of itself. $\mathbf{C}_k$ is a matrix of size $N\times N$ representing $N(=2^k)$ codes, each of length N bits and the subscript k in $\mathbf{C}_k$ represents the layer of recursion. As shown below, a set of 2^k codes can be generated at the k^{th} layer using the recursive relation in with an initial condition

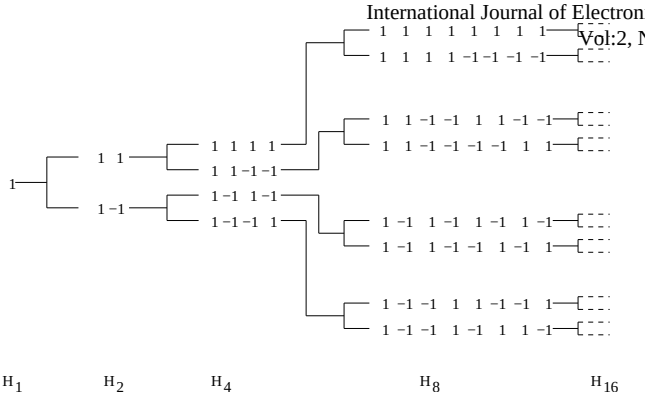


Fig. 4. OVSF code tree for WH codes.

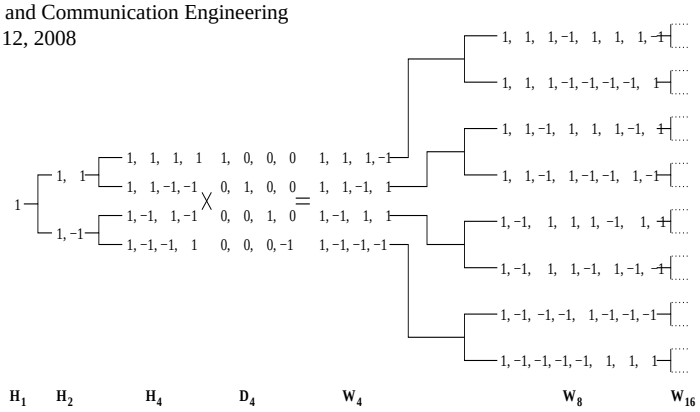


Fig. 5. OVSF code tree for MWH codes with $\{1, 1, 1, -1\}$ as basic repetitive sequence.

$\mathbf{c}_0(1) = \{1\}$:

$$\mathbf{C}_k = \begin{bmatrix} \mathbf{c}_k(1) \\ \mathbf{c}_k(2) \\ \mathbf{c}_k(3) \\ \vdots \\ \mathbf{c}_k(2^k-1) \\ \mathbf{c}_k(2^k) \end{bmatrix} = \begin{bmatrix} \mathbf{c}_{k-1}(1)\mathbf{c}_{k-1}(1) \\ \mathbf{c}_{k-1}(1)\overline{\mathbf{c}_{k-1}(1)} \\ \mathbf{c}_{k-1}(2)\mathbf{c}_{k-1}(2) \\ \mathbf{c}_{k-1}(2)\overline{\mathbf{c}_{k-1}(2)} \\ \vdots \\ \mathbf{c}_{k-1}(2^{k-1})\mathbf{c}_{k-1}(2^{k-1}) \\ \mathbf{c}_{k-1}(2^{k-1})\overline{\mathbf{c}_{k-1}(2^{k-1})} \end{bmatrix} \quad (19)$$

where any $\overline{c}(.)$ represents the complement of $c(.)$. Note that all the codes in the tree are not orthogonal to each other. They become orthogonal only if they obey the following rules:
(a) The codes in the same layer constitute the Walsh functions and, hence, are orthogonal.
(b) Any two codes of different layers are also orthogonal except for the condition that one code is the mother code of the other.

MWH codes of given length N are generated by multiplying the Hadamard matrix with a diagonal matrix of size $N \times N$. Thus, we have 2^N different diagonal matrices with $+1$ and -1 as the diagonal elements. Every diagonal matrix can restore the fixed length orthogonal property of the MWH code set, but not the variable length orthogonal property. To acquire variable length orthogonal property for MWH code set, the diagonal elements of the diagonal matrix should be the repetitive sequence of length equal to the minimum spreading factor required in the system. Fig. 5 shows the code tree for the generation of variable length orthogonal codes with $\{1, 1, 1, -1\}$ as repetitive sequence.

C. Orthogonal Gold Codes

One can find that many cross-correlation values of Gold codes are -1. By padding one zero to the original Gold codes, it is possible to make cross-correlation values to 0 at no shift among the two sequences. In fact, $2^n + 1$ orthogonal codes can be obtained by this simple zero padding. These codes are called orthogonal Gold codes. The length of these orthogonal Gold codes is 2^n, thereby making these sequences more suitable for different applications. The correlation values of these orthogonal Gold codes are nearly equal to that of the original Gold codes.

IV. Equivalence of Pseudo-Random Codes and Orthogonal Codes

The Hadamard matrices can be easily constructed from PN sequences. Below, we state it in the form of a property and then by proving it.

Property 1: **Construction of Hadamard Matrices:** If an $(N + 1) \times (N + 1)$ array is formed whose rows are each of the PN sequences, formed by same primitive polynomial, by replacing 1's with -1's and 0's with 1's of each sequence along with adding an initial row of length N and an initial column of length $(N+1)$ with all 1's, the resultant array is a $2^n \times 2^n$ Hadamard matrix.

Proof: Any $n \times n$ real matrix H_n with all its entries as ± 1 is called a Hadamard matrix if it satisfies the relation: $H_n H_n^T = nI$, as given in (14). Following the construction of H_{2^n} as given above, it is easy to check that $H_{2^n} H_{2^n}^T = 2^n I$, where $2^n = N + 1$. ■

V. Mean Square Correlation Measures

The performance of different PN sequences is usually evaluated by mean square aperiodic auto-correlation R_{AC} (MSAAC) and mean square aperiodic cross-correlation R_{CC} (MSACC) measures. These correlation measures have been introduced by Oppermann and Vucetic [7]. If $\mathbf{c}_i(n)$ represents non-delayed version of $\mathbf{c}_k(i)$, $\mathbf{c}_j(n+\tau)$ represents the delayed version of $\mathbf{c}_k(j)$ by 'τ' units and N is the length of the sequence $\mathbf{c}_i$, then the discrete aperiodic correlation function is defined as

$$\mathbf{r}_{i,j}(\tau) = \frac{1}{N} \sum_{\tau=1-N}^{N-1} \mathbf{c}_i(n)\mathbf{c}_j(n+\tau). \quad (20)$$

The mean square aperiodic auto-correlation value for a code set containing M sequences is given by

$$R_{AC} = \frac{1}{M} \sum_{i=1}^{M} \sum_{\tau=1-N, \tau \neq 0}^{N-1} |\mathbf{r}_{i,i}(\tau)|^2 \quad (21)$$

and a similar measure for the mean square aperiodic cross-correlation value is given by

$$R_{CC} = \frac{1}{M(M-1)} \sum_{i=1}^{M} \sum_{j=1, j\neq i}^{M} \sum_{\tau=1-N}^{N-1} |\mathbf{r}_{i,j}(\tau)|^2. \quad (22)$$

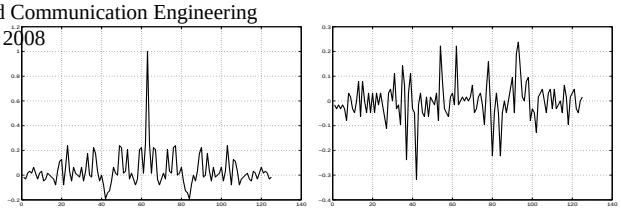
Auto-correlation refers to the degree of correspondence between a sequence and phase shifted replica of itself, whereas cross-correlation is the measure of agreement between two different codes. These two measures have been used as the basis for comparing the sequence sets in this paper. The sequences which have good auto-correlation properties will have poorer cross-correlation properties, and vice-versa, and they have wide and flat frequency spectrum. The sequences which have less MSAAC values removes the correlation among the bits with in a sample, and the sequences which have less MSACC values removes the sample to sample correlation, and make the speech signal less intelligible. We illustrate this fact by plotting aperiodic auto- and cross-correlation functions of several discussed sequences. Fig. 6(a)-(b) shows the aperiodic auto- and cross-correlation functions of Gold sequences of length 63 bits. Fig. 7(a)-(b), Fig. 8(a)-(b), Fig. 9(a)-(b) and Fig. 10 (a)-(b) show the same functions for Barker-like, Kasami, MWH and orthogonal Gold sequences, respectively. Note that for all the PN sequences, the length of the codes have been taken as 63 bits, whereas for orthogonal codes, it has been taken as 64 bits. The figures clearly show the MSAAC and MSACC tradeoff for these codes.

A. *Figure of Merit*

As has been mentioned, the price for being able to select good cross-correlation properties will be a degradation in the auto-correlation properties of the set of sequences. A degradation of the auto-correlation properties has a direct relation on the frequency spectrum of the sequences in the set. If the R_{AC} values are poor, the spectrum of the sequence will not be wide-band and flat. In order to determine quantitatively how significant this degradation is for a given set of sequences, a Figure of Merit (FoM) is required to judge the suitability of the frequency characteristics of the sequences. Sequences with low FoM has narrow flat spectrum and they are neither suitable for CDMA nor for speech encryption. The FoM for a sequence, $\mathbf{c}_i(n)$, of length N having the auto-correlation function $\mathbf{r}_i(\tau)$ is given as: $F_x = \frac{\mathbf{r}_{i,i}^2(0)}{\sum_{\tau\neq 0} |\mathbf{r}_{i,i}(\tau)|^2} = \frac{N^2}{2.\sum_{\tau=1}^{N-1} |\mathbf{r}_{i,i}(\tau)|^2}$. This is nothing more than the inverse of the MSAAC value for a given sequence. It should be noted that for the new class of PN code, given in next section, this FoM may be extended to the whole sequence set as each sequence in the set has the same absolute value of the auto-correlation. Thus we may use the inverse of the value calculated for the R_{AC} as the FoM.

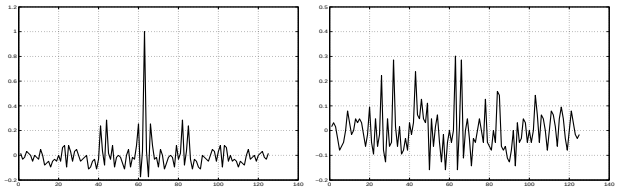
VI. Proposal of a New Class of PN Code

We construct a new family of sequences $\mathbf{e}_i = (e_i, e_{i+1}, ...)$ by generating two m-sequences $\mathbf{c}_i = (c_i, c_{i+1}, ..., c_{i+2^{n_1}-2})$ and $\mathbf{d}_i = (d_i, d_{i+1}, ..., d_{i+2^{n_2}-2})$ with two different primitive polynomials whose respective degrees n_1 and n_2 are two primes and then having a mod 2 (i) addition and (ii) multiplication of these two m-sequences. The resultant family of sequences thus generated can broadly be classified as modulo-2 operation of Mersenne (MOM) sequences. In particular, there will be two different kind of sequences: mod 2 addition



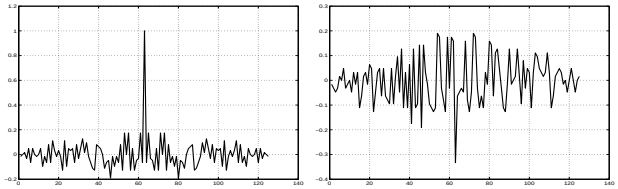
(a) (b)

Fig. 6. Aperiodic (a) auto-correlation function, (b) cross-correlation function, of Gold sequence of length 63 bits.



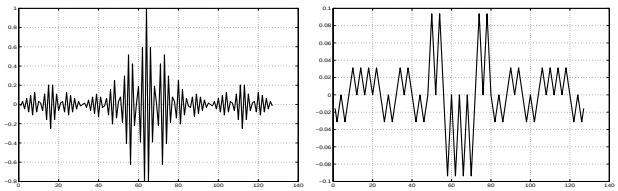
(a) (b)

Fig. 7. Aperiodic (a) auto-correlation function, (b) cross-correlation function, of Barker-like sequence of length 63 bits.



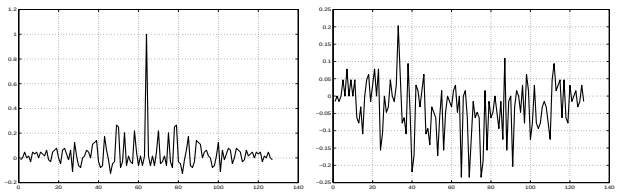
(a) (b)

Fig. 8. Aperiodic (a) auto-correlation function, (b) cross-correlation function, of large Kasami sequence of length 63 bits.



(b)

Fig. 9. Aperiodic (a) auto-correlation function, (b) cross-correlation function, of MWH sequence of length 64 bits.



(a) (b)

Fig. 10. Aperiodic (a) auto-correlation function, (b) cross-correlation function, of orthogonal Gold sequence of length 64 bits.

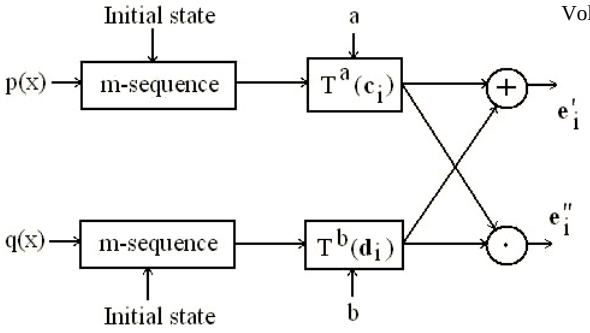


Fig. 11. The construction of MAM ($\mathbf{e}'_i$) and MMM ($\mathbf{e}''_i$) sequences.

of Mersenne (MAM) $\mathbf{e}'_i = \mathbf{c}_i + \mathbf{d}_i$ and mod 2 multiplication of Mersenne (MMM) $\mathbf{e}''_i = \mathbf{c}_i.\mathbf{d}_i$ sequences, which are shown in Fig. 11. We provide an interesting property of this new family of sequences in the sequel. Note that whenever we mention MOM in the sequel, unless otherwise specified, it implies either MAM or MMM.

Notations: We consider $\mathbf{c}_i$ is recursively generated via primitive polynomial $p(x) \in GF(2^{n_1})$ and $\mathbf{d}_i$ via $q(x) \in GF(2^{n_2})$. A k shift on the resultant recursive sequence $\mathbf{e}_i$ would be denoted as $T^k(\mathbf{e}_i)$, i.e., $T^0(\mathbf{e}_i) = \mathbf{e}_i$. If any $T^a(\mathbf{c}_i)(+/.)T^b(\mathbf{d}_i) = \tilde{\mathbf{e}}_i$ and if $\tilde{\mathbf{e}}_i = T^k(\mathbf{e}_i)$ for any k, then their shift equivalence would be denoted as $\tilde{\mathbf{e}}_i \sim \mathbf{e}_i$.

Property 2: The values of cross-correlation $R_c^{e_i,T^k(e_i)}$ between any two shift equivalent MOM sequences is bounded by the set $\{t_e, -f(n_1), -f(n_2)\}$ where any $f(n) = 2^n - 1$.

Proof: We represent the sequence as $R_c^{e_i,T^k(e_i)} = \sum_{i=0}^{M_{n_1}M_{n_2}-1}(-1)^{e_i+T^k(e_i)}$. Therefore, either in $\mathbf{e}'_i$ or in $\mathbf{e}''_i$, the cross correlation between any two sequences is equal to the difference of total number of agreements and disagreements of the two corresponding bits in any i-th position within a single period $M_{n_1}M_{n_2}$. It is known that the correlation of any length N m-sequence is defined as a set $\{N, u\}$ where N denotes the auto-correlation with zero delay while $u = -1$ represents cross-correlation (R_c) levels (the value comes from balance and window properties of m-sequences). The correlation values of MAM sequences can thus be defined as $\{M_{n_1}, -1\}\{M_{n_2}, -1\} = \{M_{n_1}M_{n_2}, 1, -M_{n_1}, -M_{n_2}\}$, where the first argument denotes auto-correlation value with zero delay (by another property we haven't shown). Hence, the cross-correlation of MAM sequences would be given by the set $R_c = \{1, -M_{n_1}, -M_{n_2}\}$. Note that in R_c, the first argument $u.u = 1$ denotes the difference between total number of zeros and total number of ones for $\mathbf{e}'_i$ and hence can be replaced by $t_{e''}$ for MMM sequences. This completes the structure of R_c for MOM sequences in general. ■

Corollary 1: The theoretical Welch lower bound on maximum level of R_c comes approximately as $\sqrt{M_{n_1}M_{n_2}}$ in our case as the number of sequences are many. From the set R_c, we get $|R_{c_{max}}| = M_{n_2} > \sqrt{M_{n_1}M_{n_2}}$. The MOM sequences thus follow the Welch lower bound criteria. On the other hand, the R_c values are considerably less in comparison with the long sequence period $M_{n_1}M_{n_2}$.

VII. Conclusion

We have reviewed different PN as well as fixed- and variable-length orthogonal sequences that can be used in many applications including spreading codes for CDMA cellular networks. The equivalence of PN and orthogonal codes are also derived. Simulation results show that large Kasami sequence has both good correlation values and high FoM, which make these sequences to have wide flat spectrum that is better suited for any applications. For this reason, these are used in the reverse link of W-CDMA systems. A new PN sequence has been proposed in the latter part of this paper which is shown to have a bounded cross-correlation value, better than the existing codes. It is expected that this new code would be explored in CDMA application areas.

References

[1] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*, Addison-Wesley, 1995.
[2] W. Diffe and M. E. Hellman, "New directions in cryptography," *IEEE Trans. Inform. Theory*, vol. 22, pp. 644-654, Nov. 1976.
[3] R. L. Pickholtz, D. L. Schilling and L. B. Milstein, "Theory of spread spectrum communications- A tutorial," *IEEE Trans. Commun.*, vol. COM-30, no. 5, May 1982.
[4] E. H. Dinan and B. Jabbari, "Spreading codes for direct sequence CDMA and wideband CDMA cellular networks," *IEEE Commun. Magazine*, vol. 36, no. 4, pp. 48-54, Sep. 1998.
[5] B. Sklar, *Digital Communications: Fundamentals and Applications*, 2^{nd} Ed., Prentice Hall, 2001.
[6] J. H. Lindholm, "An analysis of the pseudo randomness properties of the subsequences of long m-sequences," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 569-576, July 1968.
[7] I. Oppermann and B. S. Vucetic, "Complex spreading sequences with a wide range of correlation properties," *IEEE Trans. Commun.*, vol. COM-45, pp. 365-375, March 1997.
[8] L. T. Wang and E. J. McCluskey, "Linear feedback shift register design using cyclic codes," *IEEE Trans. Comput.*, vol. 37, pp. 1302-1306, Oct. 1988.
[9] A. Fuster and L. J. Garcia, "An efficient algorithm to generate binary sequences for cryptographic purposes," *Theoretical Computer Science*, vol. 259, pp. 679-688, May 2001.
[10] D. V. Sarwate and M. B. Pursley, "Correlation properties of pseudo random and related sequences," *Proc. IEEE*, vol. 68, no. 5, pp. 593-619, May 1980.
[11] S. W. Golomb and R. A. Scholtz, "Generalized Barker sequences," *IEEE Trans. Inform. Theory*, vol. IT-11, no. 4, pp. 533-537, Oct. 1965.
[12] R. Turyn and J. E. Storer, "On binary sequences," *Proc. Am. Math. Soc,* vol. 12, pp. 394-399, June 1961.
[13] D. G. Luenberger, "On Barker codes of even length," *Proc. IEEE,* vol. 51, pp. 230-231, Jan. 1963.
[14] C. K. Chan and W. H. Lam, "Generalised Barker-like PN sequences for quasisynchronous spread spectrum multiple access communication systems," *IEE Proc. Commun.*, vol. 142, no. 2, pp. 91-98, April 1995.
[15] X. Wang, Y. Wu and B. Caron, "Transmitter identification using embedded pseudo random sequences," *IEEE Tran. Broadcasting*, vol. 50, no. 3, pp. 244-252, Sep. 2004.
[16] V. Milosevic, V. Delic and V. Senk, "Hadamard transform application in speech scrambling," *Proc. IEEE*, vol. 1, pp. 361-364, July 1997.
[17] Tai-Kuo Woo, "Orthogonal variable spreading codes for wideband CDMA," *IEEE Trans. Vehicular Techn.*, vol. 51, no. 4, pp. 700-709, July 2002.
[18] E. J. Watson, "Primitive Polynomials (mod 2)," *Mathematics of Computation*, vol. 16, pp. 368-369, 1962.
[19] B. Wysocki and T. A. Wysocki, "Modified Walsh Hadamard sequences for DS-CDMA wireless systems," *School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Australia.* [Online] Available: www.elec.uow.edu.au/staff/wysocki/publications/J1.pdf.