Several Classes of Binary Sequences with Three-Level Autocorrelation

Cunsheng Ding, Tor Helleseth, *Fellow, IEEE,* and Kwok Yan Lam

***Abstract*—In this correspondence we describe several classes of binary sequences with three-level autocorrelation. Those classes of binary sequences are based on cyclic almost difference sets. Some classes of binary sequences have optimum autocorrelation.**

***Index Terms*—Almost difference set, cyclotomy, sequence.**

I. Introduction

Let D be a subset of Z_N. The characteristic sequence s^∞ of D is defined as

$$s_i = \begin{cases} 1, & \text{if } i \bmod N \in D \\ 0, & \text{otherwise.} \end{cases}$$

Let s^∞ and t^∞ be binary sequences of period N (not necessarily the least period). The periodic crosscorrelation function of the two sequences s^∞ and t^∞ is defined by

$$C_{s,t}(w) = \sum_{i \in Z_N} (-1)^{s_{i+w} - t_i}$$

where Z_N denotes the ring $\{0, 1, \cdots, N-1\}$ with integer multiplication modulo N and integer addition modulo N.

The autocorrelation function of s^∞ is defined as

$$C_s(w) = \sum_{i \in Z_N} (-1)^{s_{i+w} - s_i}.$$

Pseudorandom sequences have wide applications in simulation, software testing, global positioning systems, ranging systems, code-division multiple-access systems, radar systems, spread-spectrum communication systems, and stream ciphers. Many applications require binary sequences that have good autocorrelation properties [3], [5], [8], [10], [11], [15].

Let s^∞ be a binary sequence of period N (not necessarily the least period), and let $C = \{0 \le i \le N-1 : s_i = 1\}$. The set C is called the *characteristic set* of the sequence s^∞. The autocorrelation property of s^∞ is determined by the difference function defined as

$$d_C(w) = |(w + C) \cap C|.$$

Lemma 1 [3, p. 143]: Let s^∞ be the same as before. Then

$$C_s(w) = N - 4(k - d_C(w))$$

where $k = |C|$.

Let D be a subset of Z_N, and let $k = |D|$. D is called an (N, k, λ) difference set of Z_N if the equation $x - y = w$ has λ solutions $(x, y) \in D \times D$ for each nonzero element of Z_N. By Lemma 1 the sequence s^∞ has a two-level autocorrelation function if and only if its characteristic set C is a difference set. Thus finding binary sequences with a two-level autocorrelation function is the same as searching for difference sets of Z_N.

Clearly, in many cases Z_N has no difference sets. For instance, Z_N has no $(N, (N-1)/2, \lambda)$ difference sets if $N \equiv 1 \pmod 4$. Thus as far as autocorrelation property is concerned, in such cases we wish to get binary sequences with three-level autocorrelation. By Lemma 1 a binary sequence s^∞ has three-level autocorrelation if and only if the difference function $d_C(w)$ is three-valued.

Let D be a subset of an Abelian group $(G, +)$ such that $N = |G|$, where N is odd. D is called an (N, k, λ) almost difference set (see [6] and [3, p. 140]), if for some $(N-1)/2$ nonzero elements $a \in Z_N$, the equation

$$x - y = a$$

has exactly λ solutions $(x, y) \in D \times D$; and for the rest of $(N-1)/2$ nonzero elements there are exactly $\lambda + 1$ solutions. In other words, D is an (N, k, λ) almost difference set if and only if the difference function $d_D(w)$ takes on the value λ for half of the nonzero elements w of Z_N, and $\lambda+1$ for the other half. The (N, k, λ) almost difference sets introduced here are different from the $(m, n, k, \lambda_1, \lambda_2)$ almost difference sets introduced in [4] by Davis, but they are more or less in the same sense.

The following lemma follows directly from Lemma 1 and the definition of almost difference sets.

Lemma 2: Let C be an (N, k, λ) almost difference set of Z_N and the characteristic set of a binary sequence s^∞, i.e., $s_i = 1$ if and only if $i \bmod N \in C$. Then

$$C_s(w) = \begin{cases} N, & w = 0 \\ N - 4(k - \lambda), & \text{for half of these } w \text{ of } Z_N^* \\ N - 4(k - \lambda - 1), & \text{for the other half.} \end{cases}$$

Thus each (N, k, λ) almost difference set of Z_N gives a binary sequence with three-level autocorrelation. Of special interest are the $(N, (N-1)/2, (N-5)/4)$ almost difference sets which gives binary sequences of period N with optimum balance among 0's and 1's and with optimum autocorrelation, where $N \equiv 1 \pmod 4$.

In this correspondence, we present several classes of binary sequences with three-level autocorrelation. They are based on cyclic almost difference sets of $(Z_N, +)$, and some of them have optimum autocorrelation and optimum balance among 0's and 1's.

II. Almost Difference Sets of Z_N and Their Sequences

From the definition of (N, k, λ) almost difference sets of Z_N, it follows immediately that the following necessary condition:

$$k(k-1) = (2\lambda + 1)(N-1)/2 \tag{1}$$

holds for all (N, k, λ) almost difference sets of Z_N. It is obvious that every odd integer (≥ 3) must be of one of the two forms $4t+1$ and $4t-1$ for some t. If $N = 4t-1$ for some t, then $(N-1)/2 = 2t-1$ is odd, it follows that $(2\lambda+1)(N-1)/2$ must be odd. Thus if Z_N has an (N, k, λ) almost difference set, then N must be of the form $4t+1$.

For any subset A of Z_N and $a \in A$, we define $a + A$ to be $\{a + x : x \in A\}$, and aA to be $\{ax : x \in A\}$. Similar to difference sets [1], almost difference sets have the following basic properties.

Manuscript received July 20, 1998. This work was supported by the Norwegian Research Council under Grant Number 127203/410 and the Singapore NSTB Research Grant RP960668.

C. Ding and K. Y. Lam are with the Department of Computer Science, National University of Singapore, Singapore 119260 (e-mail: {dingcs}{lamky}@comp.nus.edu.sg).

T. Helleseth is with the Department of Informatics, University of Bergen, N-5020 Bergen, Norway (e-mail: tor.helleseth@ii.uib.no).

Communicated by R. M. Roth, Associate Editor for Coding Theory.

Publisher Item Identifier S 0018-9448(99)07446-5.

Theorem 1: Let D be an (N,k,λ) almost difference set of Z_N. Then

1) aD is also an (N,k,λ) almost difference set of Z_N if $\gcd(a,N)=1$;
2) D^* is an $(N,N-k,N-2k+\lambda)$ almost difference set of Z_N, where D^* is defined to be $D^*=Z_N\setminus D$ and is called the complement of D.

Proof: The first part of this theorem is easy to see. We prove the second part. Define

$$d_D(w)=|D\cap(D+w)|.$$

It is not difficult to see that

$$\begin{aligned}|(-w+D)\cap D^*| &= k-d_D(w)\\ |(-w+D^*)\cap D| &= k-d_D(w)\\ |(-w+D^*)\cap D^*| &= N-2k+d_D(w).\end{aligned}$$

The conclusion of the second part then follows. □

To search for almost difference sets of Z_N, we need the help of cyclotomic numbers. Let $N=df+1$ be an odd prime and let θ be a fixed primitive element of Z_N. Denote the multiplicative subgroup (θ^d) as D_0, then the coset decomposition of Z_N^* with respect to the subgroup D_0 is

$$Z_N^*=\cup_{i=0}^{d-1}D_i,$$

where $D_i=\theta^iD_0$ for $0\le i\le d-1$. The coset D_l is called the *index class* l [1] or *cyclotomic class* l [16]. Let $(l,m)_d$ denote the number of solutions (x,y) of the equation

$$1=y-x,\qquad (x,y)\in D_l\times D_m$$

or, equivalently,

$$(l,m)_d=|(D_l+1)\cap D_m|.$$

These constants $(l,m)_d$ are called *cyclotomic numbers*. Clearly, there are at most d^2 distinct cyclotomic numbers of order d and these numbers depend not only on N, d, l, and m, but also on which of the $\phi(N-1)$ primitive elements of Z_N is chosen. Cyclotomic numbers were introduced by Gauss [9], when he studied higher reciprocity, cyclotomic equations, the constructibility of regular polygons, and the quadratic partition of the form $3t+1$ into x^2+27y^2. They were used to study the Waring's problem by Dickson [2]. We now use them to search for almost difference sets.

It is known that if $N=4t+1$ is a prime, then the quadratic residues modulo N form an $(N,(N-1)/2,(N-5)/4)$ almost difference set, which can be proved easily. For biquadratic residues we have the following result.

Theorem 2 ([3], [7, p. 151]): Let a prime $N=4f+1=x^2+4y^2$ with $x\equiv 1 \pmod 4$. If f is odd, then the biquadratic residues modulo N form an $(N,f,(f-3)/4)$ almost difference set if and only if $x=5$ or -3. If f is even, they cannot form an almost difference set.

Another class of almost difference sets is described by the following theorem.

Theorem 3: Let $N=4f+1=x^2+4y^2$ with $x\equiv 1 \pmod 4$, and let D_i's be the cyclotomic classes defined before. Then $D_0\cup\{0\}$ is an almost difference set if and only if $f\equiv 1 \pmod 4$ and $x=1$ or $x=-7$.

TABLE I
THE RELATIONS OF THE CYCLOTOMIC NUMBERS OF ORDER 4, f ODD

(h,k)	0	1	2	3
0	A	B	C	D
1	E	E	D	B
2	A	E	A	E
3	E	D	B	E

TABLE II
THE RELATIONS OF THE CYCLOTOMIC NUMBERS OF ORDER 4, f EVEN

(h,k)	0	1	2	3
0	A	B	C	D
1	B	D	E	E
2	C	E	C	E
3	D	E	E	B

Proof: We consider the cyclotomic numbers of order 4. Since $N\equiv 1 \pmod 4$, N can be expressed as $N=x^2+4y^2$, $x\equiv 1 \pmod 4$, here y is two-valued, depending on the choice of the primitive root. Let D_i be the cyclotomic classes defined before.

When f is odd, the relation between the 16 cyclotomic numbers is given by Table I [2], [16].

Thus there are five possible different cyclotomic numbers in the case f being odd; i.e.,

$$\begin{aligned}A &= \frac{N-7+2x}{16}\\ B &= \frac{N+1+2x-8y}{16}\\ C &= \frac{N+1-6x}{16}\\ D &= \frac{N+1+2x+8y}{16}\\ E &= \frac{N-3-2x}{16}.\end{aligned}$$

When f is even, the relation between the 16 cyclotomic numbers is given by Table II [2], [16].

Thus there are five possible different cyclotomic numbers in the case f being even; i.e.,

$$\begin{aligned}A &= \frac{N-11-6x}{16}\\ B &= \frac{N-3+2x+8y}{16}\\ C &= \frac{N-3+2x}{16}\\ D &= \frac{N-3+2x-8y}{16}\\ E &= \frac{N+1-2x}{16}.\end{aligned}$$

Note that $|D_0\cup\{0\}|=f+1$. If $D_0\cup\{0\}$ is an almost difference set, then

$$\frac{N-1}{2}(2\lambda+1)=(f+1)f$$

which gives $\lambda=(f-1)/4$. Hence $f\equiv 1 \pmod 4$, which is odd.

We need only to consider

$$\Delta_i=:|(D_0\cup\{0\}+\theta^i)\cap(D_0\cup\{0\})| \tag{2}$$

for $i=0,1,2,3$. Note that

$$\begin{aligned}\Delta_i &= |(D_{4-i}\cup\{0\}+1)\cap(D_{4-i}\cup\{0\}|\\ &= |(D_{4-i}+1)\cap D_{4-i}|+|\{1\}\cap D_{4-i}|\\ &\quad +|(D_{4-i}+1)\cap\{0\}|\\ &= (4-i,4-i)+|\{1\}\cap D_{4-i}|+|(D_{4-i}+1)\cap\{0\}|.\end{aligned}$$

Since f is odd, it follows from

$$0 = \theta^{4f} - 1 = (\theta^{2f} - 1)(\theta^{2f} + 1)$$

that $-1 = \theta^{2f} = \theta^{8\lambda+2} \in D_2$, where θ is the primitive root of N used to define D_i. Thus (2) takes on only the following values:

$$(0,0) + 1 = A + 1 = (2,2) + 1 \qquad (1,1) = E = (3,3).$$

Hence, $D_0 \cup \{0\}$ is an almost difference set if and only if $A+1-E = \pm 1$, which are equivalent to $x = 1$ and $x = -7$, respectively. □

Note that the binary sequences based on the above two classes of almost difference sets do not have optimum balance among the 0's and 1's. We now describe a class of $(N, (N-1)/2, (N-5)/4)$ almost difference sets which give binary sequences with optimum balance of 0's and 1's and with optimum autocorrelation. By Lemma 2 the sequences induced by the almost difference sets in the following Theorems 4 and 5 have the following three autocorrelation values:

$$C_s(w) = \begin{cases} N, & w = 0 \\ -3, & \text{for half of these } w \text{ of } Z_N^* \\ 1, & \text{for the other half.} \end{cases}$$

Thus they have optimum autocorrelation.

Theorem 4: Let $N = 4f + 1 = x^2 + 4y^2$ with $x \equiv 1 \pmod 4$. Let D_i's be the cyclotomic classes of order four defined before. Then $D_0 \cup D_1$ is an $(N, (N-1)/2, (N-5)/4)$ almost difference set if and only if f is odd and $y = \pm 1$.

Proof: As before, we need only to consider

$$\begin{aligned} \Delta_i &:= |(D_0 \cup D_1 + \theta^i) \cap (D_0 \cup D_1)| \\ &= |(D_0 + \theta^i) \cap D_0| + |(D_0 + \theta^i) \cap D_1| \\ &\quad + |(D_1 + \theta^i) \cap D_1| + |(D_1 + \theta^i) \cap D_0| \\ &= (-i,-i) + (-i,-i+1) + (-i+1,-i+1 + (-i+1,-i). \end{aligned}$$

Suppose that f is odd. By the cyclotomic numbers of order 4 described before, we have

$$\Delta_0 = \Delta_2 = \frac{4N - 12 - 8y}{16}$$
$$\Delta_1 = \Delta_3 = \frac{4N - 12 + 8y}{16}.$$

Thus in this case $D_0 \cup D_1$ is an almost difference set if and only if $\Delta_1 - \Delta_0 = y = \pm 1$.

Now suppose that f is even. By the cyclotomic numbers of order 4 we have

$$\Delta_0 = A + B + D + B$$
$$\Delta_1 = B + D + A + D$$
$$\Delta_2 = C + E + B + E$$
$$\Delta_3 = D + E + C + E.$$

Note that B is not equal to D. Then $D_0 \cup D_1$ is an almost difference set if and only if

$$\Delta_0 = \Delta_2, \quad \Delta_1 = \Delta_3, \quad \Delta_0 - \Delta_1 = \pm 1$$

or

$$\Delta_0 = \Delta_3, \quad \Delta_1 = \Delta_2, \quad \Delta_0 - \Delta_1 = \pm 1.$$

It is easily checked that none of them has a solution. □

TABLE III
THE RELATIONS OF THE CYCLOTOMIC NUMBERS OF ORDER 6

(h,k)	0	1	2	3	4	5
0	(0,0)	(0,1)	(0,2)	(0,3)	(0,4)	(0,5)
1	(0,1)	(0,5)	(1,2)	(1,3)	(1,4)	(1,2)
2	(0,2)	(1,2)	(0,4)	(1,4)	(2,4)	(1,3)
3	(0,3)	(1,3)	(1,4)	(0,3)	(1,3)	(1,4)
4	(0,4)	(1,4)	(2,4)	(1,3)	(0,2)	(1,2)
5	(0,5)	(1,2)	(1,3)	(1,4)	(1,2)	(0,1)

Example 1: Let $N = 5^2 + 4 = 29$. By Theorem 4

$$D_0 \cup D_1 = \{1,2,3,7,11,14,16,17,19,20,21,23,24,25\}$$

is a $(29, 14, 6)$ almost difference set of Z_{29}. The corresponding binary sequence is

$$s^\infty = \underline{01110001000100101101110111000}\cdots$$

which is a binary sequence of period 29 with optimum autocorrelation and optimum balance between 0's and 1's.

Note that

$$D_1 \cup D_2 = \theta(D_0 \cup D_1)$$
$$D_2 \cup D_3 = \theta^2(D_0 \cup D_1)$$
$$D_3 \cup D_0 = \theta^3(D_0 \cup D_1)$$
$$D_1 \cup D_3 = \theta(D_0 \cup D_2).$$

The proof of Theorem 4 has also proved the following result.

Theorem 5: Let $N = 4f + 1 = x^2 + 4y^2$ with $x \equiv 1 \pmod 4$. Let D_i's be the cyclotomic classes of order four defined before. Then $D_1 \cup D_2$, or $D_2 \cup D_3$, or $D_3 \cup D_0$, is an $(N, (N-1)/2, (N-5)/4)$ almost difference set if and only if f is odd and $y = \pm 1$.

Let $p = 6f + 1$, and let $D_0 = (\theta^6)$ be the set of sixth powers with respect to p. By (1), a necessary condition for D_0 to be a (p, f, λ) almost difference set is that

$$f \equiv 4 \pmod 6 \quad \text{and} \quad \lambda = (f-4)/6.$$

Unfortunately, D_0 cannot be an almost difference set, as proved in the following theorem.

Theorem 6: Let $p = 6f + 1$ and $f \equiv 4 \pmod 6$. Then D_0 cannot be a $(p, f, (f-4)/6)$ almost difference set.

Proof: Note that $|D_0 \cap (D_0 + x)|$ is a constant for x in each cyclotomic class D_i. So we need only to consider $|D_0 \cap (D_0 + \theta^i)|$ for $i = 0, 1, \cdots, 5$. By definition we have as before

$$|D_0 \cap (D_0 + \theta^i)| = (6 - i, 6 - i).$$

Thus D_0 is an almost difference set if and only if among the six cyclotomic constants (i,i), $i = 0, 1, \cdots, 5$, three of them are equal to $\lambda = (f-4)/6$, and the other three equal to $\lambda + 1$.

To prove this theorem, we need the above six cyclotomic constants. It has been proven that, the 36 cyclotomic constants (k,h) depend solely upon the decomposition $A^2 + 3B^2$ of the prime $p = 6f + 1$ [2], [17]. In the case f even, there are three sets of cyclotomic numbers, depending on the choice of the primitive element θ of Z_p. Specifically, there are ten possible distinct cyclotomic numbers. The relations of these numbers are given in Table III.

The values of the ten basic constants are expressible in terms of p, A, B, and depend on the cubic character of 2 modulo p. Select

TABLE IV
THE CYCLOTOMIC NUMBERS OF ORDER 6 FOR EVEN f

	$m \equiv 0 \pmod 3$	$m \equiv 1 \pmod 3$	$m \equiv 2 \pmod 3$
36(0,0)	$p-17-20A$	$p-17-8A+6B$	$p-17-8A-6B$
36(0,1)	$p-5+4A+18B$	$p-5+4A+12B$	$p-5+4A+6B$
36(0,2)	$p-5+4A+6B$	$p-5+4A-6B$	$p-5-8A$
36(0,3)	$p-5+4A$	$p-5+4A-6B$	$p-5+4A+6B$
36(0,4)	$p-5+4A-6B$	$p-5-8A$	$p-5+4A+6B$
36(0,5)	$p-5+4A-18B$	$p-5+4A-6B$	$p-5+4A-12B$
36(1,2)	$p+1-2A$	$p+1-2A-6B$	$p+1-2A+6B$
36(1,3)	$p+1-2A$	$p+1-2A-6B$	$p+1-2A-12B$
36(1,4)	$p+1-2A$	$p+1-2A+12B$	$p+1-2A+6B$
36(2,4)	$p+1-2A$	$p+1+10A+6B$	$p+1+10A-6B$

the integer m so that $\theta^m \equiv 2 \pmod p$, then the three sets of cyclotomic numbers are given in Table IV.

By Table III, we have

$$(1,1)=(0,5), \quad (2,2)=(0,4), \quad (3,3)=(0,3),$$
$$(4,4)=(0,2), \quad (5,5)=(0,1).$$

Now we consider the six cyclotomic numbers $(0,i)$ according to the three cases. When $m \equiv 0 \pmod 3$, by Table VI, the six cyclotomic numbers $(0,i)$ take on at least four different values, so D_0 cannot form an almost difference set. When $m \equiv 1 \pmod 3$, we have

$$36(0,2)=36(0,3)=36(0,5)=p-5+4A-6B.$$

Thus if D_0 is an almost difference set, then

$$p-17-8A+6B=p-5+4A+12B=p-5-8A$$

which has the only solution $A=-2$, $B=2$. This gives $p=16$, a contradiction to the primality of p. When $m \equiv 2 \pmod 3$, we can similarly prove that the six cyclotomic numbers $(0,i)$ take on at least three different values. □

As mentioned earlier we are much interested in $(N,(N-1)/2,(N-5)/4)$ almost difference sets, as they give binary sequences with three-level autocorrelation and optimum balance among 0's and 1's. One natural question is whether there are $(N,(N-1)/2,(N-5)/4)$ almost difference sets of form $D_i \cup D_j \cup D_k$, where D_i are cyclotomic classes of order 6 and i, j, and k are pairwise-distinct. If $D_i \cup D_j \cup D_k$ is an $(N,(N-1)/2,(N-5)/4)$ almost difference set, then the necessary condition

$$(N-1)/2 \times (N-3)/2 = (N-1)/2 \times (2\lambda+1)$$

says that f must be even.

Theorem 7: Let $N=6f+1$ with f even, and let θ be a primitive root of N, which is used to define the cyclotomic classes of order 6. Assume that $m \equiv 1 \pmod 3$, where $\theta^m=2$. Then $D_0 \cup D_1 \cup D_2$ is an $(N,(N-1)/2,(N-5)/4)$ almost difference set if and only if $N=13$ or $N=61$.

Proof: Define

$$\Delta(a)=|(D_0 \cup D_1 \cup D_2 + a) \cap (D_0 \cup D_1 \cup D_2)|$$

where $a \in Z_N^*$. With the cyclotomic numbers of order 6 described before, it is computed that $\Delta(a)$ takes on the following six values:

$$\Delta_0=\frac{9p-45}{36}$$
$$\Delta_1=\frac{9p-21-12B}{36}$$
$$\Delta_2=\frac{9p-15-6A+6B}{36}$$
$$\Delta_3=\frac{9p-9}{36}$$
$$\Delta_4=\frac{9p-33+12B}{36}$$
$$\Delta_5=\frac{9p-39+6A-6B}{36}.$$

When $A=7$ and $B=2$, we obtain

$$\Delta_0=\Delta_1=\Delta_2=(9p-45)/36$$

and

$$\Delta_3=\Delta_4=\Delta_5=(9p-9)/36=\Delta_0+1.$$

When $A=-2$ and $B=-1$, we have

$$\Delta_1=\Delta_2=\Delta_3=(9p-9)/36$$

and

$$\Delta_0=\Delta_4=\Delta_5=(9p-45)/36=\Delta_1-1.$$

Thus the two cases give such almost difference sets. It is checked that only the two cases lead to such almost difference sets. They correspond to $p=7^2+3\times 2^2=61$ and $p=(-2)^2+3\times(-1)^2=13$. Since both 13 and 61 have primitive root 2, the corresponding m in the two cases is 1. □

Example 2: Let $N=13$. Then by Theorem 7

$$D_0 \cup D_1 \cup D_2 = \{1,2,4,9,11,12\}$$

is a $(13,6,2)$ almost difference set. The corresponding binary sequence is

$$s^\infty = \underline{0110100001011}\cdots.$$

The following two theorems can be similarly proved as Theorem 7.

Theorem 8: Let $N=13$. Then 2 is a primitive root of N. Let 2 be the primitive root used to define the cyclotomic classes of order 6. Then $D_0 \cup D_2 \cup D_3$ is an $(N,(N-1)/2,(N-5)/4)$ almost difference set.

Theorem 9: Let $N=73$. Then $D_0 \cup D_3 \cup D_4$ is an $(N,(N-1)/2,(N-5)/4)$ almost difference set, where D_i are the cyclotomic classes of order 6 with respect to 73.

Let $N=8t+1$. It is possible for the set of octic residues $D_0=(\alpha^8)$ to form an almost difference set of Z_N, where α is a primitive root of N. Since $|D_0|=t$, a necessary condition for D_0 to be an almost difference set is $t(t-1)=(2\lambda+1)(N-1)/2$. It follows that $t=8\lambda+5$ and, therefore,

$$N=8t+1=64\lambda+41=16(4\lambda+2)+9.$$

Under these necessary conditions the cyclotomic numbers of order 8 are given in two sets of formulas according to whether 2 is a quartic residue or not, in terms of N, x, y, a, and b which are determined by [13]

$$N=x^2+4y^2=a^2+2b^2 \qquad (x \equiv a \equiv 1 \pmod 4). \quad (3)$$

For the case 2 is a quartic residue the following result is known.

Theorem 10 ([7, p. 55], [3, p. 152]): Let $N=8t+1$ and $t=8\lambda+5$, where λ is a positive integer. Assume that 2 is a quartic residue modulo N. Then the set of octic residues D_0 forms an almost difference set if and only if N admits the simultaneous representations

$$N=19^2+4y^2=1+2b^2$$

or

$$N=13^2+4y^2=1+2b^2.$$

For the case 2 is not a quartic residue the following result is known.

Theorem 11 ([7, p. 55], [3, p. 152]): Let $N = 8t + 1$ and $t = 8\lambda + 5$, where λ is a positive integer such that 2 is not a quartic residue. Then the set of octic residues D_0 forms an almost difference set if and only if $N = 41$.

The linear span (linear complexity) of a sequence is defined to be the length of the shortest linear feedback shift register that produces the sequence [8], [11]. The linear span of all the sequences defined by the almost difference sets presented before can be computed. For example, we prove the following result.

Theorem 12: Let s^∞ be the sequence with characteristic set $D_0 \cup D_1$ defined before, where D_i are cyclotomic classes of order 4. If $2 \in D_0$, then

$$L(s^\infty) = (N-1)/2.$$

If $2 \notin D_0$, then $L(s^\infty) = N - 1$, where $L(s^\infty)$ denotes the linear span (also called linear complexity).

Proof: Define

$$S^N(x) = s_0 + s_1 x + \cdots + s_{N-1}x^{N-1}.$$

It is well known [8] that the linear complexity of s^∞ is given by

$$N - \deg\left(\gcd\left(x^N - 1, S^N(x)\right)\right). \tag{4}$$

Let β be a primitive Nth root of unity over the field GF(2^m) that is the splitting field of $x^N - 1$. Then by (4) we have

$$L(s^\infty) = N - |\{j : S(\beta^j) = 0,\ 0 \le j \le N-1\}|$$

where $S(x)$ is defined by

$$S(x) = \sum_{i \in D_0 \cup D_1} x^i.$$

Define

$$T(\beta) = \sum_{i \in D_1 \cup D_2} \beta^i.$$

By definition, $aD_i = D_{i+j}$ if $a \in D_j$. Note that

$$\left(\sum_{i\in D_0} + \sum_{i\in D_1} + \sum_{i\in D_2} + \sum_{i\in D_3}\right)\beta^i = 1.$$

It follows that

$$S(\beta^d) = \begin{cases} \sum_{i\in D_0\cup D_1} \beta^i = S(\beta), & d \in D_0 \\ \sum_{i\in D_1\cup D_2} \beta^i = T(\beta), & d \in D_1 \\ \sum_{i\in D_2\cup D_3} \beta^i = S(\beta)+1, & d \in D_2 \\ \sum_{i\in D_0\cup D_3} \beta^i = T(\beta)+1, & d \in D_3. \end{cases} \tag{5}$$

Also we have

$$S(1) = 0. \tag{6}$$

We first consider the case $2 \in D_0$. Note that $2D_i = D_i$, we have

$$\begin{aligned} (S(\beta))^2 &= S(\beta^2) \\ &= \sum_{d\in D_0\cup D_1} \beta^{2d} \\ &= \sum_{d\in 2D_0\cup 2D_1} \beta^{d} \\ &= \sum_{d\in D_0\cup D_1} \beta^{d} \\ &= S(\beta). \end{aligned}$$

Hence $S(\beta) \in \{0, 1\}$. Similarly, we have $T(\beta) \in \{0, 1\}$.

Independent of whether $S(\beta)$ and $T(\beta)$ take on 1 or 0, by (5) we have

$$|\{j : S(\beta^j) = 0, 1 \le j \le N-1\}| = 2f.$$

Whence

$$L(s^\infty) = N - 1 - \frac{N-1}{2} = \frac{N-1}{2}.$$

This proves the first part of this theorem.

When $2 \in D_1$, we obtain that

$$\begin{aligned} S(\beta^2) &= S(\beta)^2 = T(\beta) \\ T(\beta^2) &= T(\beta)^2 = S(\beta) + 1. \end{aligned}$$

It follows that $S(\beta) \notin \{0, 1\}$ and $T(\beta) \notin \{0, 1\}$.

When $2 \in D_2$, we obtain that

$$\begin{aligned} S(\beta^2) &= S(\beta)^2 = S(\beta) + 1 \\ T(\beta^2) &= T(\beta)^2 = T(\beta) + 1. \end{aligned}$$

It follows that $S(\beta) \notin \{0, 1\}$ and $T(\beta) \notin \{0, 1\}$.

When $2 \in D_3$, we obtain that

$$\begin{aligned} S(\beta^2) &= S(\beta)^2 = T(\beta) + 1 \\ T(\beta^2) &= T(\beta)^2 = S(\beta). \end{aligned}$$

It follows that $S(\beta) \notin \{0, 1\}$ and $T(\beta) \notin \{0, 1\}$.

Thus when $2 \notin D_0$ we have that

$$S(\beta) \notin \{0, 1\} \quad \text{and} \quad T(\beta) \notin \{0, 1\}.$$

It then follows from (5) and (6) that

$$L(s^\infty) = N - 1. \qquad \square$$

Theorem 12 shows that the sequence with characteristic set $D_0 \cup D_1$ has good linear span.

III. Concluding Remarks

In this correspondence, we have presented several classes of almost difference sets of Z_N. Those $(N, (N-1)/2, (N-5)/4)$ almost difference sets give binary sequences of period N with optimum autocorrelation and optimum balance between 0's and 1's. They have also good linear span.

As mentioned earlier, finding binary sequences with some three-level autocorrelation values is equivalent to finding almost difference sets of Z_N with corresponding parameters. It turns out that finding almost difference sets is as hard as finding difference sets. Cyclotomy is a helpful tool in finding both difference sets and almost difference sets. However, it is quite limited. It is possible to construct almost difference sets of Z_N with cyclotomic classes of order $2e$, where $e \ge 4$. We have tried this for cyclotomic classes of order 8, but were unable to obtain any $(N, (N-1)/2, (N-5)/2)$ almost difference sets.

It would be interesting to point out whether the almost difference sets in this correspondence are related to difference sets and partial difference sets. Since we are interested only in cyclic almost difference sets for the constructions of sequences, we will mention the connections only under the context of the almost difference sets of Z_N.

As pointed out in Section II, if Z_N has an almost difference set, then $N \equiv 1 \pmod 4$. If $N \equiv 3 \pmod 4$, then Z_N could have difference sets, but not almost difference sets. If $N \equiv 1 \pmod 4$, then Z_N may have both difference sets and almost difference sets. Certain difference sets with special parameters can be used to construct almost difference sets, and *vice versa*. Details about these

TABLE V
KNOWN CYCLOTOMIC ALMOST DIFFERENCE SETS OF Z_N

cyclotomic ADS	conditions	references
$D_0^{(2,N)}$	$N \equiv 1 \pmod 4$	Paley PDS
$D_0^{(4,N)}$	$N = x^2 + 4y^2$, $x = 5$ or $x = -3$	[7, 3]
$D_0^{(4,N)} \cup \{0\}$	$N = x^2 + 4y^2$, $x = 1$ or $x = -7$	this paper
$D_i^{(4,N)} \cup D_{i+1}^{(4,N)}$	$N = x^2 + 4$, $x \equiv 1 \pmod 4$	this paper
$D_0^{(8,N)}$	$N = 64t + 41$ and $N = 19^2 + 4y^2 = 1 + 2b^2$ or $N = 13^2 + 4y^2 = 1 + 2b^2$	[7, 3]
$D_0^{(6,13)} \cup D_1^{(6,13)} \cup D_2^{(6,13)}$		this paper
$D_0^{(6,61)} \cup D_1^{(6,61)} \cup D_2^{(6,61)}$		this paper
$D_0^{(6,13)} \cup D_2^{(6,13)} \cup D_3^{(6,13)}$		this paper
$D_0^{(6,73)} \cup D_3^{(6,73)} \cup D_4^{(6,73)}$		this paper

TABLE VI
KNOWN CYCLOTOMIC DIFFERENCE SETS OF Z_N

cyclotomic DS	conditions	references
$D_0^{(2,N)}$	$N \equiv 3 \pmod 4$	Paley DS
$D_0^{(4,N)}$	$N = 4t^2 + 1$, t odd	[12]
$D_0^{(4,N)} \cup \{0\}$	$N = 4t^2 + 9$, t odd	[12]
$D_0^{(8,N)}$	$N = 8t^2 + 1 = 64u^2 + 9$, where t and u odd	[12]
$D_0^{(8,N)} \cup \{0\}$	$N = 8t^2 + 49 = 64u^2 + 441$, where t odd, u even	[12]
$D_0^{(6,N)} \cup D_1^{(6,N)} \cup D_3^{(6,N)}$	$N = 4t^2 + 27$, $N \equiv 1 \pmod 6$	Hall DS [12]

connections will be given in a future paper by Arasu and the first two coauthors of this correspondence.

Cyclotomic classes can be used to construct both difference sets and almost difference sets of Z_N, where N is a prime. All the almost difference sets described in this correspondence are cyclotomic. It would be interesting to make a comparison between the cyclotomic difference sets and cyclotomic almost difference sets. This is done by summarizing them in Tables V and VI, where N is a prime, ADS stands for almost difference sets, DS denotes difference sets, and PDS stands for partial difference sets (definition will be given below). The two tables illustrate the connections and differences.

Let G be an Abelian group of order v and D a subset of G with $|D| = k$. Then D is called a (v, k, λ, μ)-partial difference sets if for every nonidentity element g of D, the equation $d_1 - d_2 = g$ has exactly λ solutions $(d_1, d_2) \in D \times D$; and for every nonidentity element g' of $G \setminus D$, the equation $d_1 - d_2 = g'$ has exactly μ solutions [14]. Here we are concerned with only Abelian partial difference sets.

The (v, k, λ) almost difference sets and $(v, k, \lambda, \lambda + 1)$ partial difference sets are in general quite different. For the former we have less restriction on g and more restriction on the number of elements g such that $d_1 - d_2 = g$ has λ solutions, while for the latter we have more restriction on g and less restriction on the number of elements g such that $d_1 - d_2 = g$ has λ solutions. Among all the known cyclotomic almost difference sets in Table V there is one that is also a partial difference set, as stated in the following theorem.

Theorem 13: When $N \equiv 1 \pmod 4$ is a prime, the set $D_0^{(2,N)}$ of quadratic residues modulo N is both an almost difference set and a partial difference set, called the Paley partial difference set.

Proof: It is very easy to prove the two conclusions by using cyclotomic numbers of order 2 [3], [16]. □

For some applications (e.g., stream ciphering), binary sequences with good balance between the number of 0's and that of 1's may be better. However, in other applications it may not be necessary to require a balance between them. So almost difference sets D of Z_N with $|D|$ being not far away from $N/2$ could also have important applications. On the other hand, in the definition of almost difference sets the condition that $d_1 - d_2 = g$ has λ solutions for half of the nonzero elements may be weakened and such sets could give sequences with good autocorrelation.

ACKNOWLEDGMENT

The authors wish to thank the referee for his detailed and constructive comments and suggestions that considerably improved this correspondence.

REFERENCES

[1] L. D. Baumert, *Cyclic Difference Sets* (Lecture Notes in Mathematics, vol. 182). New York: Springer-Verlag, 1971.
[2] L. E. Dickson, "Cyclotomy, higher congruences, and Waring's problem," *Amer. J. Math.*, vol. 57, pp. 391–424, and 463–474, 1935.
[3] T. W. Cusick, C. Ding, and A. Renvall, *Stream Ciphers and Number Theory* (North-Holland Mathematical Library, vol. 55). Amsterdam, The Netherlands: North-Holland/Elsevier, 1998.
[4] J. A. Davis, "Almost difference sets and reversible difference sets," *Arch. Math.*, vol. 59, pp. 595–602, 1992.
[5] C. Ding, "The differential cryptanalysis and design of the natural stream ciphers," in *Fast Software Encryption*, R. Anderson, Ed. (Lecture Notes in Computer Science, vol. 809). Heidelberg, Germany: Springer-Verlag, 1994, pp. 101–115.
[6] ——, "Binary cyclotomic generators," in *Fast Software Encryption*, B. Preneel, Ed. (Lecture Notes in Computer Science, vol. 1008). New York: Springer-Verlag, 1995, pp. 29–60.
[7] ——, *Cryptographic Counter Generators* (TUCS Series in Dissertation, no. 4), Turku Centre for Computer Science, ISBN 951-650-929-0, 1997.

[8] C. Ding, G. Xiao, and W. Shan, *The Stability Theory of Stream Ciphers* (Lecture Notes in Computer Science, vol. 561). Heidelberg, Germany: Springer-Verlag, 1991.
[9] C. F. Gauss, *Disquisitiones Arithmeticae*, Leipzig, Germay, 1801; English translation: New Haven, CT, Yale Univ., 1966; reprint by Springer-Verlag, Berlin, Heidelberg, and New York, 1986.
[10] S. W. Golomb, *Shift-Register Sequences.* San Francisco, CA: Holden-Day, 1967; Laguna Hills, CA: Aegean Park, 1982.
[11] T. Helleseth and P. V. Kumar, "Sequences with low correlation," in *Handbook of Coding Theory*, V. Pless and W. C. Huffman, Eds. Amsterdam, The Netherlands: Elsevier, 1998.
[12] D. Jungnickel and A. Pott, "Difference sets: Abelian," in *The CRC Handbook of Combinatorial Designs*, C. J. Colbourn and J. H. Dinitz, Eds. New York: CRC, 1996, pp. 297–307.
[13] E. Lehmer, "On the number of solutions of $u^2 + D \equiv w^2 \pmod p$," *Pacific J. Math.*, vol. 5, pp. 103–118, 1955.
[14] S. L. Ma, "A survey of partial difference sets," *Des., Codes Cryptogr.*, vol. 4, pp. 221–261, 1994.
[15] D. V. Sarwate, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, vol. 68, pp. 593–619, 1980.
[16] T. Storer, *Cyclotomy and Difference Sets.* Chicago, IL: Markham, 1967.
[17] A. L. Whiteman, "The cyclotomic numbers of order twelve," *Acta Arith.*, vol. 6, pp. 53–76, 1960.

Fast Coding of Low-Entropy Sources

Boris Ya. Ryabko and Marina P. Sharova

***Abstract*— The problem of coding low-entropy information sources is considered. Since the run-length code was offered about 50 years ago by Shannon, it is known that for such sources there exist coding methods much simpler than for sources of a general type. However, known coding methods of low-entropy sources do not reach the given redundancy. In this correspondence, a new method of coding low-entropy sources is offered. It permits a given redundancy r with almost the same encoder and decoder memory size as that obtained by Ryabko for general methods, while encoding and decoding much faster.**

***Index Terms*— Complexity of coding, fast algorithm, low-entropy sources, redundancy, run-length coding.**

I. Introduction

We consider the problem of low-entropy source coding whose elementary example is a Bernoulli source generating a sequence of zeros and ones with probabilities q and p, respectively, when $p \to 0$. This problem has attracted attention of many researchers, as for coding of such sources there exist simpler methods than in a general case. The efficiency of a code is measured by redundancy and by complexity of encoding and decoding. The redundancy r is the difference between the average codeword length and the Shannon entropy. Complexity is estimated by the memory size of the encoder and decoder (in bits) and by the average time of encoding and decoding one symbol measured by the number of binary operations on single-bit word when they are implemented on a computer with random-access memory (see the definition in [2]).

One of the well-known compression schemes of low-entropy sources is run-length coding [1]. In this method, a sequence of symbols generated by a source is broken into runs of zeros between two sequential ones: 1, 01, 001, etc., then the lengths of the runs are encoded by the binary codewords. The length of a run can thus be both limited and unlimited.

In coding with unlimited length of runs the scheme offered by Shannon [1] can be used. According to this scheme, one codeword is selected for the least probable symbol 1. For encoding lengths of runs binary words are picked in ascending order, bypassing the word selected for 1. Shannon has proved that by increasing the length of the codeword designating 1 when $p \to 0$ the redundancy of coding tends to zero. It is possible to show that it does not exceed $C_1 p \log(1/p)$, where $C_1 \leq 1$ is a constant.

In [3] Elias proposes to use a prefix code of integers for run-length coding. Elias has constructed three new universal binary representations of integers and by using them has constructed universal codeword sets. For the best representation of integers from [3] the redundancy of it the given code reaches $C_2 p \log\log(1/p)$, where C_2 is a constant.

An effective run-length coding method was offered by Golomb [4]. In [5] it was shown that for particular values of a run-length coding scheme Golomb's code is optimal.

However, the known methods of coding low-entropy sources [1], [3]–[5] do not allow reaching the given redundancy. In this correspondence, a new method of coding low-entropy sources is offered. It permits reaching a given redundancy r with almost the same encoder and decoder memory size as obtained in [6] for general methods, while encoding and decoding is much faster.

Here we consider a problem of coding a Bernoulli source with known statistics. Note that the offered code construction is applicable also for the Bernoulli sources with unknown statistics and for more complex models.

II. Algorithm of Coding Low-Entropy Sources

Let a Bernoulli source generating a sequence of zeros and ones with probabilities q and p, respectively, when $p \to 0$, be given. Let $r > 0$ be the given redundancy of a code. Our problem is to construct a method of source coding permitting us to reach the given redundancy r.

In our method encoding is implemented in two stages: first, a message is compressed by a simple code and an output sequence is then encoded by a fast and effective code. After the first stage the length of the input sequence is essentially reduced, and applying a complex fast algorithm at the second stage provides little total time of encoding and decoding per letter of the initial message. At the second stage it is possible to use many codes, for example, the arithmetic code [7], [8] or the code from [6]. We shall use the code from [6] since it has the estimates of the average time and memory. Note, however, that the use of some versions of the universal arithmetic code gives the same result. For code from [6] the dependence of the memory size V and the average time T of encoding and decoding of one letter on the redundancy r' as $r' \to 0$ satisfies the following

Manuscript received February 1, 1998; revised January 14, 1999. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Cambridge, MA, August 1998.

B. Ya. Ryabko is with the Siberian State Academy of Telecommunications and Computer Science, 630102 Novosibirsk, Russia.

M. P. Sharova is with the Novosibirsk State University, 630102 Novosibirsk, Russia.

Communicated by I. Csiszár, Associate Editor for Shannon Theory.
Publisher Item Identifier S 0018-9448(99)07675-0.