Filomat 27:3 (2013), 459–466
DOI 10.2298/FIL1303459O

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Lifted Polynomials Over F_{16} and Their Applications to DNA Codes

Elif Segah Oztas, Irfan Siap

Yildiz Technical University, Department of Mathematics, Istanbul, Turkey

Abstract. In this paper, we introduce a new family of polynomials which generates reversible codes over a finite field with sixteen elements (F_{16} or $GF(16)$). We name the polynomials in this family as *lifted polynomials*. Some advantages of lifted polynomials are that they are easy to construct, there are plenty of examples of them and it is easy to determine the dimension of codes generated by them. Furthermore we introduce *4-lifted polynomials* which provide a rich source for DNA codes. Also we construct codes over F_4 that have the best possible parameters from lifted polynomials. In addition we obtain some reversible codes over F_4.

1. Introduction

The interest on DNA computing started by the pioneer paper written by Leonard Adleman [3]. Adleman solved a hard (NP- complete) computational problem by DNA molecules in a test tube. DNA sequences consist of four bases (nucleotides) that are (A) adenine, (G) guanine, (T) thymine and (C) cytosine. DNA has two strands that are arranged in an order with a rule that is named as Watson Crick complement (WCC). Briefly the WCC of A is T and vice versa and the WCC of G is C and vice versa. We describe this as $A^c = T, T^c = A, G^c = C$ and $C^c = G$.

In [2], constructions of DNA codes over the finite field with four elements $GF(4)$ are presented. Examples that have larger sizes comparing to the previous examples in the literature are constructed therein. Also, [9, 10, 12] focused on constructing large sets of DNA codewords. In [19], DNA codes over a four element ring $F_2[u]/\langle u^2 - 1\rangle$ are considered. In [21], for the first time instead of single DNA bases, double DNA pairs are matched to a 16 element ring and the algebraic structure of these DNA codes are studied. Hamming distance constraint, reverse-complement constraint, reverse constraint, fixed GC-content constraint are the most common constraints used in DNA codes [9, 12, 13, 15]. In [20], stochastic search algorithms are used to design codewords for DNA computing by Tuplan et al. [5] and [6] present genetic and evolutionary algorithms from sets of DNA sequences. In [18], cyclic codes over $GF(4)$ are used to construct DNA codes. But, they constrict their study to only linear reversible cyclic codes over $GF(4)$. In [7] and [8] deletion similarity distance is used that is different from Hamming distance and more suitable for DNA codes. In [14], Mansuripur et al. show that DNA molecules can be used as a storage media. In [1], a new approach is developed that extends the work in [10] and [12] and reverse-complement constraint is added to further prevent unwanted hybridizations. Studies about DNA computing indicate that DNA computing will be

2010 *Mathematics Subject Classification*. Primary 92D10; Secondary 11T71, 94B05.
Keywords. Reversible codes; Lifted polynomials; DNA codes.
Received: 29 November 2012; Accepted: 24 February 2012
Communicated by Eberhard Malkowsky
Email addresses: elifsegahoztas@gmail.com (Elif Segah Oztas), isiap@yildiz.edu.tr (Irfan Siap)

of much interest in the near future. Hence, DNA Error Correcting Codes will be of great value since DNA computing is faster and can store more memory than silicon based computing systems.

Reversible codes are useful for DNA structure. Reversible codes were introduced by Massey over $GF(q)$ (q is a prime power) in [16]. In [17], Muttoo and Lal have studied reversible codes over $GF(q)$ (q is prime). But, these codes are obtained from a specially constructed parity check matrix. In [4], Das and Tyagi give the generalized form of parity check matrix to obtain reversible codes both odd and even length.

In this paper, we use a special family of polynomials to construct reversible codes different from [16] and [17]. The polynomials are called "lifted polynomials" and "4-lifted polynomials" which are introduced by authors and these polynomials satisfy the reversibility over F_{16}. By Means of these new families of polynomials and the construction presented, it is possible to construct DNA codes where some properties can be controlled. Further, by applying this approach we not only obtain odd length DNA codes which is the case in [21], but also we obtain even length DNA codes. Also, we provide some examples of linear codes with best possible parameters from lifted polynomial construction which can be considered as DNA codes.

The rest of paper is organized as follows. In Section 2, we give some basic notions. In Section 3, we introduce reversible codes obtained by lifted polynomial. In Section 4 we introduce DNA codes that are generated by using 4-lifted polynomial over F_{16}. In Section 5, there are some applications about the codes that have the best possible parameters over F_4. Section 6 concludes the paper.

2. Background

Let F be a finite field. A linear code of length n over F is an F-vector space of F^n. A cyclic code C of length n over F is invariant with respect to the right cyclic shift operator that maps a codeword $(c_0, c_1, ..., c_{n-1}) \in C$ to another codeword $(c_{n-1}, c_0, ..., c_{n-2})$ in C. For each codeword $(c_0, c_1, ..., c_{n-1})$, we associate the polynomial $g(x) = c_0 + c_1x + ... + c_{n-1}x^{n-1}$ where $c_i \in F$. Let

$$\begin{aligned} \Phi :& F[x]/(x^n - 1) \to C \\ & g(x) = c_0 + c_1x + ... + c_{n-1}x^{n-1} \to (c_0, c_1, ..., c_{n-1}) \end{aligned} \tag{1}$$

For each codeword $c = (c_0, c_1, ..., c_{n-1})$, we define the reverse of c to be $c^r = (c_{n-1}, c_{n-2}, ..., c_0)$.

Definition 2.1. *A linear code C of length n over F is said to be reversible if $c^r \in C$ for all $c \in C$.*

The Hamming distance between codewords c' and c'', denoted by $H(c', c'')$, is simply the number of coordinates in which these two codewords differ.

For each polynomial $z(x) = z_0 + z_1x + ... + z_rx^r$ with $z_r \neq 0$, the reciprocal of $z(x)$ is defined to be the polynomial $z^*(x) = x^r z(1/x) = z_r + z_{r-1}x^{r-1} + ... + z_0x^r$. Consider that $\deg z^*(x) \leq \deg z(x)$ and if $z_0 \neq 0$, then $z(x)$ and $z^*(x)$ always have the same degrees. $z(x)$ is called self-reciprocal if and only if $z(x) = z^*(x)$.

Theorem 2.2. *[16] The cyclic code generated by a monic polynomial $g(x)$ is reversible if and only if $g(x)$ is self-reciprocal where $g(x)|(x^n - 1)$.*

3. Reversible codes over F_{16} obtained by lifted polynomials

Here, we define lifted polynomials which generate reversible codes under the restriction that the length n is odd.

Definition 3.1. *Let $g(x) = a_0 + a_1x + ... + a_tx^t$ be a self reciprocal polynomial over Z_2 and $g(x)|(x^n - 1)$ $(mod\ 2)$. A lifted polynomial of $g(x)$ is denoted by $\ell_g(x) \in F_{16}[x]$ and is defined as follows. If t is odd, then*

$$\ell_g(x) = \sum_{i=0}^{\frac{t-1}{2}} \theta_i; \quad \theta_i = \begin{cases} \beta_i x^i + \beta_i x^{t-i} & , a_i \neq 0 \\ 0 & , a_i = 0 \end{cases}, \tag{2}$$

if t is even, then

$$\ell_g(x) = \sum_{i=0}^{\frac{t}{2}} \theta_i; \quad \theta_i = \begin{cases} \beta_i x^i + \beta_i x^{t-i} & , a_i \neq 0\ ,\ i \neq t/2 \\ 0 & , a_i = 0 \\ \beta_{\frac{t}{2}} x^{\frac{t}{2}} & , a_i \neq 0\ ,\ i = t/2 \end{cases} \tag{3}$$

where $\beta_i \in F_{16} - \{0\}$. There are many lifted polynomials of $g(x)$ depending on β_i. The set of $\ell_g(x)$ is denoted by $\mathcal{L}_g(x)$. The terms x^i and x^{t-i} are called complement pairs.

Example 3.2. *Here we present some examples of lifted polynomials. Let $n = 15$ and*

$$g(x) = 1 + x + x^3 + x^4 + x^5 + x^7 + x^8$$

be a self reciprocal polynomial and where $g(x)|(x^{15} - 1)$. We can write some arbitrary lifted polynomials of $g(x)$ as follows,

$$\ell_g(x) = \alpha^{10} + \alpha x + \alpha^2 x^3 + \alpha^5 x^4 + \alpha^2 x^5 + \alpha x^7 + \alpha^{10} x^8$$

and

$$\ell_g(x) = \alpha^5 + x + \alpha^6 x^3 + \alpha^2 x^4 + \alpha^6 x^5 + x^7 + \alpha^5 x^8$$

where $\alpha^i \in F_{16}$, $i \in \{0, 1, ..., 14\}$.

Lemma 3.3. *Lifted polynomials are self reciprocal polynomials over F_{16}.*

Proof. The proof follows easily by observing that the lifted polynomials are obtained by self reciprocal polynomials over Z_2 by Definition 3.1. □

Lemma 3.4. *Suppose that a set S consists of vectors and their reverses. Then, the code generated by S as an F-vector subspace is reversible.*

Proof. Suppose that every element in S has its reverse in S. Let S be a spanning set and

$$S = \langle c_1, c_1^r, c_2, c_2^r, ..., c_k, c_k^r, c_{s_1}, c_{s_2}, ..., c_{s_m} \rangle$$

where c_i is codeword and $c_{s_t} = c_{s_t}^r$ (self reversible) where $1 \leq t \leq m$. Let $C = \langle S \rangle$. For every codeword $c = \sum \beta_i c_i \in C$, since $(\alpha c_i + \beta c_j)^r = \alpha c_i^r + \beta c_j^r$, we have $c^r = \sum \beta_i c_i^r \in C$ where $i, j \in \{1, 2, ..., k, s_1, ..., s_m\}$. Hence C is a reversible code. □

Remark 3.5. *In this paper, the notation $\langle S \rangle$ will denote the F-vector space generated by the set S. The notation (S) will stand for the ideal generated by S.*

Theorem 3.6. *Let $\ell_g(x)$ be a lifted polynomial over F_{16} of a self reciprocal polynomial $g(x)$ and $g(x)|(x^n - 1)(mod 2)$ with $\deg(g(x)) = t$ where n is odd. Let*

$$S = \{\ell_g(x), x\ell_g(x), ..., x^{n-t-1}\ell_g(x)\} \tag{4}$$

and $C_l = \langle S \rangle$. If C_l is a code which is generated by the spanning set S, then C_l is a reversible F_{16}-linear code and it is shortly denoted by $C_l = \langle \ell_g(x) \rangle$.

Proof. By Lemma 3.4 the reverse of each codeword can be found as follows:

$$\left(\Phi\left(\sum_{j=1}^{n-t}\beta_{i_j}x^{i_j}\ell_g(x)\right)\right)^r=\Phi\left(\sum_{j=1}^{n-t}\beta_{i_j}x^{n-t-1-i_j}\ell_g(x)\right) \tag{5}$$

since $0\le i_j\le n-t-1$ we have $n-t-1\ge n-t-1-i_j\ge 0$, where $\beta_{i_j}\in F_{16}$, $i_j\in\{0,1,...,n-t-1\}$ and Φ is given in (1). The claim follows. □

Note that $\langle\ell_g(x)\rangle$ is not an ideal but an F_{16}-linear code spanned by S.
We now present an example of a reversible code, by making use of lifted polynomials.

Example 3.7. *Let $n=15$ and $g(x)=1+x+x^3+x^4+x^5+x^7+x^8$ be a self reciprocal polynomial and $g(x)|(x^{15}-1)$ and $d=3$ for $C=(g(x))$. We can choose a lifted polynomial $\ell_g(x)=\alpha^{10}+\alpha x+\alpha^2x^3+\alpha^5x^4+\alpha^2x^5+\alpha x^7+\alpha^{10}x^8$.*
We know $k=7$ (dimension of code C) then we choose this value k for $C_l=\langle\ell_g(x)\rangle$. Then, C_l gives a $[15,7,7]_{16}$ reversible code obtained over F_{16}.

In this section, all the statements are suitable for reversible codes over F_4. Further we present some example of a reversible code over $F_4=GF(4)=F_2[x]/(x^2+x+1)$ which is obtained by a lifted polynomial. This linear code has the best possible parameters.

Example 3.8. *Let $n=9$ and $g(x)=1+x+x^2$ be a self reciprocal polynomial and $g(x)|(x^9-1)$ and $d=2$ for $C=(g(x))$ over Z_2. We can choose an arbitrary $\ell_g(x)=\omega+x+\omega x^2$ over $F_4=\{0,1,\omega,\omega^2=1+\omega\}$. We know $k=7$ (dimension of code C) then we choose this value k for $C_l=\langle\ell_g(x)\rangle$. Then, C_l gives a $[9,7,2]_4$ reversible code obtained over F_4. This code attains the best possible parameters ([11]).*

Example 3.9. *Let $g(x)=1+x+x^3+x^4+x^5+x^7+x^8$ and $g|(x^{15}-1)(mod\ 2)$. Let $\ell_g(x)=1+\omega x+\omega x^3+\omega x^4+\omega x^5+\omega x^7+x^8$ over F_4. We can represent the spanning set as a matrix. Hence, the generator matrix is*

$$\begin{pmatrix}
1&w&0&w&w&w&0&w&1&0&0&0&0&0&0\\
0&1&w&0&w&w&w&0&w&1&0&0&0&0&0\\
0&0&1&w&0&w&w&w&0&w&1&0&0&0&0\\
0&0&0&1&w&0&w&w&w&0&w&1&0&0&0\\
0&0&0&0&1&w&0&w&w&w&0&w&1&0&0\\
0&0&0&0&0&1&w&0&w&w&w&0&w&1&0\\
0&0&0&0&0&0&1&w&0&w&w&w&0&w&1
\end{pmatrix}$$

and $C_l=\langle\ell_g(x)\rangle$ gives a $[15,7,7]_4$-reversible code. This code has the best possible parameters [11].

4. DNA codes over F_{16} from 4-lifted polynomial

In this section, we mainly mention about DNA codes over F_{16}. But, there is a reversibility problem for DNA codes over F_{16}. Reversibility problem: Let $(\alpha,\alpha^2,1)$ be a codeword and corresponds to ATGCTT in DNA where $\alpha\to AT,\alpha^2\to GC,1\to TT$ and $\alpha,\alpha^2,1\in F_{16}$. Reverse of $(\alpha,\alpha^2,1)$ is $(1,\alpha^2,\alpha)$. $(1,\alpha^2,\alpha)$ corresponds to TTGCAT. But, TTGCAT is not reverse of ATGCTT. Because, reverse of ATGCTT is TTCGTA. We have solved this problem with 4-lifted polynomial that is introduced by authors.

DNA occur in sequences, represented by sequences of letters from the alphabet $S_{D_4}=\{A,T,G,C\}$. We define a DNA code of length $2n$ to be a set of codewords $(\alpha_0,...,\alpha_{n-1})$ where n is odd.

We consider $GF(16)=F_{16}=F_2[x]/\left(x^4+x+1\right)$ for DNA and DNA double bases (pairs).

$$\alpha_i\in\{AA,AT,AG,AC,TT,TA,TG,TC,GG,GA,GC,GT,CC,CA,CG,CT\}=S_{D_{16}}. \tag{6}$$

The most difficult and interesting problem is to provide a matching between the field elements and DNA alphabets. This matching should obey the rules and properties of DNA. Here, we accomplish this task by

Table 1: 4-power table

double DNA pair	F_{16}(multiplicative)	additive
AA	0	
TT	α^0	1
AT	α^1	α
GC	α^2	α^2
AG	α^3	α^3
TA	α^4	$1+\alpha$
CC	α^5	$\alpha+\alpha^2$
AC	α^6	$\alpha^2+\alpha^3$
GT	α^7	$1+\alpha+\alpha^3$
CG	α^8	$1+\alpha^2$
CA	α^9	$\alpha+\alpha^3$
GG	α^{10}	$1+\alpha+\alpha^2$
CT	α^{11}	$\alpha+\alpha^2+\alpha^3$
GA	α^{12}	$1+\alpha+\alpha^2+\alpha^3$
TG	α^{13}	$1+\alpha^2+\alpha^3$
TC	α^{14}	$1+\alpha^3$

presenting Table 1 which is called the 4-power table that gives a correspondence between DNA and F_{16}. In this table, there is a property that the related DNA double pair of an element of F_{16} is reverse of the related DNA double pair of fourth power of the element of F_{16}. For instance, $\alpha^2 \to GC$ and $(\alpha^2)^4 = \alpha^8 \to CG$.

The Watson-Crick complement is given by $A^c = T, T^c = A, C^c = G, G^c = C$. Hence we use Watson-Crick complement such that $(AA)^c = TT, ..., (TC)^c = AG$. If α is a codeword, we show $\alpha = (\alpha_0, ..., \alpha_{n-1})$. We define the complement of α to be $\alpha^c = (\alpha_0^c, ..., \alpha_{n-1}^c)$, the reverse-complement of α to be $\alpha^{rc} = (\alpha_{n-1}^c, ..., \alpha_0^c)$.

The following definition describes how to match the codewords over the field with DNA codewords.

Definition 4.1. *Let C be a code over F_{16} of length n and $c \in C$ be a codeword where $c = (c_0, c_1, ..., c_{n-1})$, and $c_i \in F_{16}$. We define*

$$\Theta(c) : C \to S_{D_4}^{2n} \text{ where } (c_0, c_1, ..., c_{n-1}) \to (b_0, b_1, ..., b_{2n-1}) \quad (7)$$

where each of c_i is mapped to coordinate pairs (b_{2i}, b_{2i+1}) for $i = \{0, 1, ..., n-1\}$ defined in Table 1. Hence, $\Theta(c) = (b_0, b_1, ..., b_{2n-1})$ is a DNA codeword of $\Theta(C)$ where $b_j \in S_{D_4}$, for $j \in \{0, 1, ..., 2n-1\}$.

For instance, $(c_0, c_1, c_2, c_3) = (\alpha, \alpha^5, \alpha^6, \alpha^{11}) \to (ATCCACCT) = (b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7)$.

In order to obtain DNA codes, we define and use a new family of the lifted polynomials over F_{16}.

Definition 4.2. *Let $g(x)$ be a self reciprocal polynomial over Z_2 and $g(x)|(x^n-1)$ with $\deg(g(x)) = t$. A 4-lifted polynomial of $g(x)$ is denoted by $l_g^{(4)}(x) \in F_{16}[x]$*

if t is odd, then

$$\ell_g^{(4)}(x) = \sum_{i=0}^{\frac{t-1}{2}} \tau_i; \quad \tau_i = \begin{cases} \beta_i^v x^i + \beta_i^{4v} x^{t-i} & , a_i \neq 0 \\ 0 & , a_i = 0 \end{cases}, \quad (8)$$

if t is even, then

$$\ell_g^{(4)}(x) = \sum_{i=0}^{\frac{t}{2}} \tau_i; \quad \tau_i = \begin{cases} \beta_i^v x^i + \beta_i^{4v} x^{t-i} & , a_i \neq 0, i \neq t/2 \\ 0 & , a_i = 0 \\ \beta_{\frac{t}{2}} x^{\frac{t}{2}} & , a_i \neq 0, \beta_{\frac{t}{2}} \in \{0, 1, \alpha^5, \alpha^{10}\}, i = t/2 \end{cases} \quad (9)$$

where $\beta_i \in F_{16} - \{0\}$, $\ell_g^{(4)}(x)$ is shortly denoted by $\ell^{(4)}$ polynomial of $g(x)$.

In the previous section, complement pairs had the same coefficients, however in this definition one is the fourth power the other.

Example 4.3. *Let $n = 15$ and*

$$g(x) = 1 + x + x^3 + x^4 + x^6 + x^7$$

be a self reciprocal polynomial where $g(x)|(x^{15} - 1)$. Some examples of 4-lifted polynomials of $g(x)$ are as follows

$$\ell_g^{(4)}(x) = \alpha + \alpha^4 x + \alpha^7 x^3 + \alpha^{7*4} x^4 + \alpha^{4*4} x^6 + \alpha^4 x^7$$

and

$$\ell_g^{(4)}(x) = \alpha^4 + \alpha^2 x + \alpha^6 x^3 + \alpha^9 x^4 + \alpha^8 x^6 + x^7.$$

Theorem 4.4. *Let $\ell_g^{(4)}(x)$ be a 4-lifted polynomial of a self reciprocal polynomial $g(x)$ over F_{16} and $g(x)|(x^n - 1)(mod 2)$ with $\deg(g(x)) = t$ where n is odd. Let*

$$S = \{\ell_g^{(4)}(x), x\ell_g^{(4)}(x), ..., x^{n-t-1}\ell_g^{(4)}(x)\} \quad (10)$$

and $C_{\ell^{(4)}} = \langle S \rangle$. Then $\Theta(C_{\ell^{(4)}})$ is a reversible DNA code of length $2n$ which is denoted by $C_{\ell^{(4)}} = \langle \ell_g^{(4)}(x) \rangle$.

Proof. $\Theta(\sum_i x^i l_g^{(4)})$ determines DNA codewords of $\Theta(C_{\ell^{(4)}})$. Reverses of DNA codewords are denoted as follows

$$\Theta(\sum_i x^i l_g^{(4)}(x))^r = \Theta(\sum_i x^{n-t-1-i} l_g^{(4)}(x))$$

where $i \in \{0, 1, ..., n - t - 1\}$. By Lemma 3.4, $\Theta(C_{\ell^{(4)}})$ is a reversible DNA code because of the structure of the spanning set S for $C_{\ell^{(4)}}$. □

Note that $\langle \ell_g^{(4)}(x) \rangle$ is not an ideal but an F_{16}- linear code spanned by S.

Theorem 4.5. *Let $\ell_g^{(4)}(x)$ be a 4-lifted polynomial over F_{16} of a self reciprocal polynomial $g(x)$ and $g(x)|(x^n - 1)(mod 2)$ with $\deg(g(x)) = t$ where n is odd. Let $r(x) = 1 + x + ... + x^{n-1}$ and*

$$S = \{\ell_g^{(4)}(x), x\ell_g^{(4)}(x), ..., x^{n-t-1}\ell_g^{(4)}(x), r(x)\} \quad (11)$$

and $C_{\ell^{(4)}} = \langle S \rangle$. If $C_{\ell^{(4)}}$ is a code which generated by the spanning set S, then $\Theta(C_{\ell^{(4)}})$ is a reversible complement DNA code of length $2n$ and $C_{\ell^{(4)}}$ is denoted by $C_{\ell^{(4)}} = \langle \ell_g^{(4)}(x), r(x) \rangle$.

Proof. In addition to proof of Theorem 4.4, we need to consider their complements too. First we observe that for each $\alpha \in F_{16}$ we have $\alpha + 1 = \alpha^c$. In order to include the complements of the codewords, it suffices to have $r(x)$ included in the code. □

Example 4.6. *Let $C = (g(x))$ be a reversible linear code length of 9 over Z_2 where $g(x) = 1 + x + x^3 + x^4 + x^6 + x^7$ is a self reciprocal polynomial. Let $\ell_g^{(4)}(x) = \alpha + \alpha^4 x + \alpha^7 x^3 + \alpha^{7*4} x^4 + \alpha^{4*4} x^6 + \alpha^4 x^7$ be a $\ell^{(4)}$ polynomial over F_{16}. We find $(18, 256, 6)_{S_{D_4}}$ reversible DNA code that is generated by $C_{\ell^{(4)}} = \langle \ell_g^{(4)}(x) \rangle$.*

For $C_{\ell^{(4)}} = \langle \ell_g^{(4)}(x), r(x) \rangle$, $\Theta(C_{\ell^{(4)}})$ gives $(18, 4096, 5)_{S_{D_4}}$ reversible complement DNA code.

Table 2: $[9,3,5]$ reversible complement DNA code

AAAAAAAAA	TTTTTTTTT	CCCCCCCCC	GGGGGGGGG
ATCAGGACT	TAGTCCTGA	CGACTTCAG	GCTGAAGTC
ACGATTAGC	TGCTAATCG	CATCGGCTA	GTAGCCGAT
AGTACCATG	TCATGGTAC	CTGCAACGT	GACGTTGCA
TCAGGACTA	AGTCCTGAT	GACTTCAGC	CTGAAGTCG
TGCGAGCGT	ACGCTCGCA	GTATCTATG	CATAGATAC
TAGGCTCCC	ATCCGAGGG	GCTTAGAAA	CGAATCTTT
TTTGTCCAG	AAACAGGTC	GGGTGAACT	CCCACTTGA
CGATTAGCA	GCTAATCGT	ATCGGCTAC	TAGCCGATG
CCCTCGGAT	GGGAGCCTA	AAAGATTCG	TTTCTAAGC
CTGTATGTC	GACATACAG	AGTGCGTGA	TCACGCACT
CATTGCGGG	GTAACGCCC	ACGGTATTT	TGCCATAAA
GTACCATGA	CATGGTACT	TGCAACGTC	ACGTTGCAG
GACCTGTTT	CTGGACAAA	TCAAGTGGG	AGTTCACCC
GGGCGTTAC	CCCGCAATG	TTTATGGCA	AAATACCGT
GCTCACTCG	CGAGTGAGC	TAGACAGAT	ATCTGTCTA

Table 3: some examples for the codes that have the best possible parameters, obtained by lifted polynomials over F_4

n	Lifted polynomial	Code parameters
9	$1+\omega x+x^2$	[9,7,2]
9	$1+\omega^2 x+x^2$	[9,7,2]
15	$1+\omega x+x^2$	[15,13,2]
15	$1+\omega^2 x+x^2$	[15,13,2]
15	$1+x+\omega x^2+x^3+x^4$	[15,11,4]
15	$1+x+\omega^2 x^2+x^3+x^4$	[15,11,4]
15	$1+\omega x+x^2+\omega x^3+x^4$	[15,11,4]
15	$1+\omega x+\omega^2 x^2+\omega x^3+x^4$	[15,11,4]
15	$1+\omega x+\omega x^3+\omega x^4+\omega x^5+\omega x^7+x^8$	[15,7,7]
15	$1+\omega^2 x+\omega^2 x^3+\omega^2 x^4+\omega^2 x^5+\omega^2 x^7+x^8$	[15,7,7]

5. Some applications of DNA codes

We can apply lifted polynomials to generate reversible codes and reversible DNA codes over F_4 where $F_4=\{0,1,\omega,\omega^2\}$ and $A\to 0, T\to 1, C\to \omega, G\to \omega^2$. All of the statements and results of Section 3 over F_{16} also hold true for F_4 .

Example 5.1. *Let $C=(g(x))$ be a reversible linear code length of 9 over Z_2 where $g(x)=1+x+x^3+x^4+x^6+x^7$. Let $\ell_g(x)=1+\omega x+\omega^2x^3+\omega^2x^4+\omega x^6+x^7$ be a lifted polynomial over $F_4=\{0,1,\omega,\omega^2\}$. C_ℓ is a $[9,2,6]_4$ reversible DNA code. If $C_\ell=\langle \ell_g(x), r(x)\rangle$, then C_ℓ gives $[9,3,5]_4$ reversible complement DNA codes (See Table 2). In this example, we use lifted polynomial for DNA code over F_4. In this Table 2 the $[9,3,5]_4$ reversible complement DNA code for illustration purposes is presented explicitly.*

In Table 3, some codes that are generated by lifted polynomials are given. These codes have the best possible parameters over F_4. These constructions are direct constructions comparing to [11] where the codes with such parameters are obtained via indirect methods such as "shortening a code" or other. Table 3 shows that a code can be generated by different lifted polynomials. In other words, lifted polynomials provide rich families of codes. Also, these codes give reversible DNA codes over F_4.

6. Conclusion

In this paper we studied reversible codes of odd length over F_{16} and DNA codes of even length which are obtained from F_{16}. We introduce lifted and 4-lifted polynomials to construct the reversible codes and the reversible DNA codes. We found some reversible codes of odd length obtained by lifted polynomial over F_4. These codes have the best possible parameters and also they are reversible codes. We introduce a correspondence between DNA double (bases) pairs and the elements of a finite field of size 16. Hence, we introduce a new method for obtaining DNA codes of even length over F_{16}. This new method introduced here is an ongoing research of the authors. An open and interesting problem is to establish a bound for the minimum distance of this family of codes. Open problems include the study of DNA codes that are generated by lifted polynomials over F_q and their properties. Also, it will be interesting to construct dual of the codes that are generated by lifted polynomials.

Acknowledgement: We would like to express our sincere thanks to the referee for his/her valuable remarks and careful reading that helped us to improve the presentation of our work tremendously.

References

[1] N. Aboluion, D. H. Smith, Perkins S, Linear and nonlinear constructions of DNA codes with Hamming distance d, constant GC-content and a reverse-complement constraint, Discrete Mathematics 312 (2012) 1062-1075.
[2] T. Abualrub, A. Ghrayeb, X. N. Zeng, Consruction of cyclic codes over F_4 for DNA computing, Journal of the Franklin Ins., 343 (2006), 448–457.
[3] L. Adleman, Molecular computation of solutions to combinatorial problems, Science 266 (1994) 1021-1024.
[4] P.K. Das and Vinod Tyagi, Reversible and (n_1, n_2)-blockwise reversible code, International Journal of Algebra, Vol. 6 (2012) no. 4, 185–191.
[5] R. Deaton, M. Garzon, R.C. Murphy, D.R. Franceschetti, S.E. Stevens, Genetic search of reliable encoding for DNA based computation, First Conference on Genetic Programming GP-96, Stanford University, (1996) 9-15.
[6] R. Deaton, R.E. Murphy, J.A. Rose, M. Garzon, D.R. Franceschetti, S.E. Stevens Jr., A DNA based implementation of an evolutionary search for good encodings for DNA computation, IEEE International Conference on, Apr (1997) 267–272.
[7] A. Dyachkov, A. Macula, T. Renz, P. Vilenkin, I. Ismagilov, New results on DNA codes, International Symposium on Information Theory (ISIT 2005), September (2005) 283-287.
[8] A. Dyachkov, P. Erdos, A. Macula, D. Torney, C.H. Tung, P. Vilenkin , P.S. White, Exordium for DNA codes, Journal of Combinatorial Optimization, 7 (2003) 369-379.
[9] A.G. Frugtos, Q. Liu, A.J. Thiel, A.M.W. Sanner, A.E. Condon, L.M. Smith, R.M. Corn, Demonstration of a Word Design Strategy for DNA Computing on Surfaces, Nucleic Acids Res., 25 (1997) 47484757.
[10] P. Gaborit, O.D. King, Linear construction for DNA codes, Theoretical Computer Science, 334 (13) (2005) 99-113.
[11] M. Grassl, Bounds on the minimum distance of linear codes and quantum codes. Online available at http://www.codetables.de. Accessed on 2012-10-20.
[12] O. King, Bounds for DNA Codes with Constant GC-Content, Journal of Combinatorics, 10 (2003) 1-13.
[13] M. Li, H.J. Lee, A.E. Condon, R.M. Corn, DNA word design strategy for creating sets of non-interacting oligonucleotides for DNA microarrays, Langmuir 18 (2002) 805-812.
[14] M. Mansuripur, P.K. Khulbe, S.M. Kuebler, J.W. Perry, M.S. Giridhar, N. Peyghambarian, Information storage and retrieval using macromolecules as storage media, Optical Data Storage - OSA Technical Digest Series (Optical Society of America), 2003.
[15] A. Marathe, A.E. Condon, R.M. Corn, On combinatorial DNA word design, J. Comput. Biol. 8 (2001) 201-220.
[16] J.L. Massey, Reversible codes, Information and Control, 7 (1964) 369-380.
[17] S.K. Muttoo, S. Lal, A reversible code over $GF(q)$, Kybernetika, 22 N. 1 (1986) 85–91.
[18] V. Rykov, A.J. Macula, D. Torney, P. White, DNA sequences and quaternary cyclic codes, IEEE International Symposium on Information Theory (ISIT 2001), Washington, DC, (2001) p. 248.
[19] I. Siap, T. Abualrub, A. Ghrayeb, Cyclic DNA codes over the ring $F_2[u]/(u^2 - 1)$ based on the deletion distance, Journal of the Franklin Ins., 346 (2006) 731–740.
[20] D.C. Tuplan, H. Hoos, A. Condon, Stochastic Local Search Algorithms for DNA Word Design, Lecture Notes in Computer Science, Springer, Berlin, (2003) 229-241.
[21] B. Yildiz, I. Siap, Cyclic codes over $F_2[u]/(u^4 - 1)$ and applications to DNA codes, Computers and Mathematics with Application 63 (2012) 1169–1176.