ComTech: Computer, Mathematics and Engineering Applications, 10(1), June 2019, 23-27 P-ISSN: 2087-1244
DOI: 10.21512/comtech.v10i1.4716 E-ISSN: 2476-907X

Maximality on Construction of Ternary Cross Bifix Free Code

Mohammad Affaf

Department Pendidikan Matematika, STKIP PGRI Bangkalan
Jln. Soekarno Hatta No.52, Jawa Timur 69116, Indonesia
mohaffaf@stkippgri-bkl.ac.id

Received: 28th April 2018/ **Revised:** 30th January 2019/ **Accepted:** 14th February 2019

How to Cite: Affaf, M. (2019). Maximality on Construction of Ternary Cross Bifix Free Code. *ComTech: Computer, Mathematics and Engineering Applications, 10*(1), 23-27. https://doi.org/10.21512/comtech.v10i1.4716

Abstract **-** The purpose of this research was to show that ternary cross bifix free code $CBFS_3(2m+1)$ and $CBFS_3(2m+2)$ achieved the maximum for every natural number m. This research was a literature review. A cross bifix free codes was constructed by using Dyck path method which achieved the maximality, that was non-expandable on binary set sequences for appropriate length. This result is obtained by partitioning members of $CBFS_3(2m+1)$ and $CBFS_3(2m+2)$ and comparing them with the maximality of $CBFS_2(2m+1)$ and $CBFS_2(2m+2)$. For small length 3, the result also shows that the code $CBFS_3(3)$ is optimal.

Keywords: maximality, construction, ternary cross bifix free code

I. INTRODUCTION

Synchronization problem between transmitter and receiver on the data frame in the system of communication is called frame synchronization. Frame synchronization is one of the main topics in digital communication systems. In this system, to guarantee the synchronization between a transmitter and receiver, it can be done by periodically inserting a fixed sequence into the transmitted data. To find out what the transmitted data is, the receiver should find the fixed sequence. The technique is introduced by Massey (1972) and claimed that the search process is found by Nielsen (1973).

Frame synchronization method is not only used in digital communication systems but also in gene expression as shown by Weindl and Hagenauer (2007). In fixed sequences, the gene uses a fixed sequence to sign the fundamental expression first (Levy & Yaakobi, 2017). This situation gives the probability that the frame synchronization method can be used on simulation genome.

Frame synchronization method can be done by transmitting data coming from code $\{x_1, x_2\ x_3,...,x_k\}$ and having a special property. To recognize the beginning of the frame, it should guarantee that all suffixes of x_i do not occur as the prefix of x_j for all x_i, and x_j belongs to $\{x_1, x_2\ x_3,...,x_k\}$. This method is firstly introduced by Van Wijngaarden and Willink (2000). This kind of code is called as cross bifix free code/set.

At the beginning of the 21st century, the research of cross bifix free code arose to resolve frame synchronization via distributed sequence method with the generalization of bifix free studies. Cross bifix free code is a set of sequences in which no prefix of any length less than n of any sequences is the suffix of any sequence in the set. A cross bifix free code is constructed by using Dyck path method which achieves maximality and is non-expandable on binary set sequences for appropriate length.

Many researchers propose algorithms to construct the cross bifix free code since it has a wide impact on practice. Bajic (2007) was the first researcher that constructed the code by using a Kernel set method and was extended in Bajic and Loncar-Turukalo (2014). Then, Bilotta, Pergola, and Pinzani (2012) introduced a binary cross bifix free code construction for an arbitrary length by using Dyck Path. Moreover, a ternary cross bifix free code is constructed by generalizing the construction of cross bifix free codes which using Dyck path (Affaf and Ulum, 2017a, 2017b).

The code constructed using the Dyck path also extends on Bernini, Bilotta, Pinzani, Sabri, and Vajnovszki (2014) and Bilotta, Grazzini, Pergola, and Pinzani (2013). Then, Chee, Kiah, Purkayastha, and Wang (2013) introduced the construction of cross bifix free code using alphabet having q symbol. Furthermore, The construction by Chee *et al*. (2013) is generalized by Blackburn (2015). All the codes constructed are claimed to achieve the maximality, non-expandable on the appropriate set of sequences, except the code in Affaf and Ulum (2017a, 2017b). However, it is unknown whether the ternary codes on Affaf and Ulum (2017a, 2017b) achieve maximality of cross bifix free codes or not. If it is not, the construction cannot be categorized as good construction because a singleton set of bifix free sequences is also a cross bifix free code. In this research, the maximality of ternary cross bifix free code by Affaf and Ulum (2017a, 2017b) will be explained.

In this research, the researcher explains about the maximality of ternary cross bifix free code especially the maximality of $CBFS_3(2m+1)$ and $CBFS_3(2m+2)$ for every natural number m.

II. METHODS

This research is a literature review. The method used is collecting a number of information relating to construction and maximality of cross bifix free code. From this amount of information, it will be used to show that ternary cross bifix codes $CBFS_3(2m+1)$ and $CBFS_3(2m+2)$ on Affaf and Ulum (2017a, 2017b) to achieve the maximum.

The researcher briefly explains the definition and terminology on cross bifix free codes and the construction of cross bifix free code by Bilotta *et al*. (2012), Affaf and Ulum (2017a, 2017b), that is $CBFS_2(n)$, $CBFS_3(2m+1)$, and $CBFS_3(2m+2)$ respectively. The construction by Bilotta *et al*. (2012) is a cross bifix free set, $CBFS_2(n)$, which is non-expandable on $H_2(n)$. This implies that for every h belongs to $H_2(n)$, andit does not lie in $CBFS_2(n)$. Then, $CBFS_2(n)\cup\{h\}$ is not cross bifix free code. On the other side, The construction by Affaf and Ulum (2017a, 2017b) shows that $CBFS_3(2m+1)$ and $CBFS_3(2m+2)$ are a cross bifix free code. Next, the researcher will also give the construction of cross bifix free set by Chee *et al*. (2013), $S^k_{q,n}$, which is non-expandable on $H_q(n)$.

In the definitions and terminology on cross bifix free code, Σ is a finite set with cardinality q. The element of Σ is a symbol, and Σ is an alphabet. The set of all finite sequence (may be an empty sequence) on Σ denoted by Σ^* and element of Σ^* is word or codeword. Then, it will be $\Sigma^+ = \Sigma^*\backslash\{\varepsilon\}$ where ε is an empty sequence. For example, $\Sigma=\{0,1\}$, ε, 101, 00011, 1110001 is the element of Σ^*.

For ω, an element on Σ^+ is with $\omega = uvw$ where u and w are in Σ^+ and v in Σ^*. Then, u and w are prefix, and ω as suffix. Those are denoted by pre(ω) dan suf(ω) respectively. For prefix and suffix ω with length k is denoted by $pre_k(\omega)$ and $suf_k(\omega)$ respectively. From the definition of prefix and suffix, it is clear that the length of prefix or suffix of a codeword in Σ^+ is less than the length of the codeword.

A bifix of codeword ω is a word which appears as prefix dan suffix of ω. A codeword in Σ^+ is bifix free if there is no $pre_k(\omega)$ appearing as $suf_k(\omega)$. Furthermore, non-empty subset C of Σ^n, that is subset of sequence on Σ with length n, is called cross bifix free code with length n if, for every ω_i and ω_j in C, there is no $pre_k(\omega_i)$ appearing as $suf_k(\omega_j)$ for any k which is less than n. For example, for $\Sigma=\{0,1\}$, codeword 1010101 in Σ^+ contains 3 bifix codes, that are 1, 101, and 10101. Then, sets of codeword 0000111, 000110011, 0001011, 0001101, 0010101 which are subset of Σ^7 is cross bifix free code with 7 length.

Then, there are three constructions of binary cross bifix free code $CBFS_2(n)$. First, it is the construction of $CBFS_2(2m+1)$. Cross bifix free code of $CBFS_2(2m+1)$ is defined as:

$$CBFS_2\,(2m+1)=\{x\alpha : \alpha \in D_{2m}\} \qquad (1)$$

It is the set of paths beginning with a rising step linked to a $2m$-length of Dyck path. For example, for $n = 7$, the researcher shows that $CBFS_2(7)$ has elements of 1111000, 1110100, 1110010, 1101100, and 1101010. $|pre_k\,\omega|_i$ and $|suf_k\,\omega|_i$ denote the number of i which occurs in $pre_k\,\omega$ and $suf_k\,\omega$ respectively. It can be noted that for every $0<k<n$ holds $|pre_k\,\omega|_0 > |pre_k\,\omega|_1$ and $|suf_k\,\omega|_0 \leq |suf_k\,\omega|_1$ for every ω in $CBFS_2(2m+1)$. Thus, $CBFS_2(2m+1)$ is the set of bifix free sequences.

On the other side, it can also be noted that for every γ and ω in $CBFS_2(2m+1)$ holds $|pre_k\,\omega|_0 > |pre_k\,\omega|_1$ and $|suf_k\,\omega|_0 \leq |suf_k\,\omega|_1$. Now, if $CBFS_2(2m+1)$ is not cross bifix free code, there are α and β in $CBFS_2(2m+1)$ like $pre_k\alpha = suf_k\beta$ for some k which satisfies $0 < k < n$. Consequently, there is also $|pre_k\alpha|_i = |suf_k\beta|_i$ for $i \in\{0,1\}$. Then, since it is $|pre_k\alpha|_0 > |pre_k\alpha|_1$, it will be $|suf_k\beta|_0 > |suf_k\beta|_1$. However, this is a contradiction with $|suf_k\omega|_0 \leq |suf_k\omega|_1$ for every ω in $CBFS_2(2m+1)$. Thus, the set of $CBFS_2(2m+1)$ must be a cross bifix free code.

The cardinality of $CBFS_2(2m+1)$ is given by C_m, the *m-th* Catalan number, for natural number $m \geq 1$. From the construction of $CBFS_2(2m+1)$, Bilotta *et al*. (2012) got **Theorem II.B.1.1.** For $m \geq 1$, $CBFS_2(2m+1)$ was a cross bifix free code which was non-expandable on $H_2(2m+1)$.

Second, it is the construction of $CBFS_2(2m+2)$ for m even. Cross bifix free code $CBFS_2(2m+1)$ for m even is defined by Bilotta *et al*. (2012) as:

$$CBFS_2(2m+2) = \left\{\alpha x\beta\bar{x} : \alpha \in D_{2i}, \alpha \in D_{2(m-i)}, 0 \leq i \leq \frac{m}{2}\right\} \qquad (2)$$

It is the set of paths consisting of the following consecutive subpaths: a $2i$-length Dyck path, a rise step, a $2(m\text{-}i)$-length Dyck path with $0 \leq i \leq \frac{m}{2}$, and a fall step.

For example, for $n = 10$, the researcher gets that $CBFS_2(10)$ has elements of 1111100000, 1111010000, 1110110000, 1111001000, 1110101000, 1101110000, 1111000100, 1101101000, 1110100100, 1101011000, 1101100100, 1110010100, 1110011000, 1101010100, 1011110000, 1011101000, 1011100100, 1011011000, 1011010100, 1010111000, 1010110100, 1100110100, and 1100111000. Then, the cardinality of $CBFS_2(2m+2)$ for m even is given by $\sum_{i=0}^{\frac{m}{2}} C_i C_{m-i}$ for integer $m \geq 0$. From the construction of $CBFS_2(2m+2)$ for m even, Bilotta *et al*. (2012) got **TheoremII.B.2.1.** For $m \geq 0$, $CBFS_2(2m+2)$ was a cross bifix free code which was non-expandable on $H_2(2_m+2)$.

Third, it is the construction of $CBFS_2(2m+2)$ for m odd cross bifix free code $CBFS_2(2m+2)$ for m odd, defined by Bilotta *et al*. (2012) as

$$CBFS_2(2m+2) = \left\{\alpha x\beta\bar{x} : \alpha \in D_{2i}, \alpha \in D_{2(m-i)}, 0 \leq i \leq \frac{m+1}{2}\right\} \backslash \left\{x a\bar{x} x\beta\bar{x} : \alpha \in D_{2i}, \alpha \in D_{2(m-1)}\right\} \qquad (3)$$

It shows the set of paths consisting the following consecutive subpaths: a 2i-length Dyck path, a rise step, a $2(m\text{-}i)$-length Dyck path for $0 \leq i \leq \frac{m+1}{2}$, a fall step, and excluding those consisting of the following consecutive subpaths: a rise step, a $(m\text{-}1)$-length Dyck path, a fall step followed by a rise step, a $(m\text{-}1)$-length Dyck path, and a fall step. For example, for $n = 8$, the researcher obtains that $CBFS_2(8)$ has elements 11110000, 11101000, 11011000, 11100100, 11010100, 10111000, 10110100, and 10101100. In cardinality of $CBFS_2(2m+2)$ for m odd, it is given by $\sum_{i=0}^{\frac{m+1}{2}} C_i C_{m-i} - C^2_{\frac{m-1}{2}}$ for natural number $m \geq 1$. From the construction of $CBFS_2(2m+1)$ for m odd, Bilotta *et al.* (2012) obtained Theorem II.B.3.1. For $m \geq 1$, $CBFS_2(2m+2)$ was a Cross Bifix Free Set which was non-expandable on $H_2(2m+2)$.

Next, there is also the construction of ternary cross bifix free code, $CBFS_3(2m+1)$. The construction of cross bifix free code by Affaf and Ulum (2017a) is:

Construction II.C.1.1. Let $CBFS_2(2m+1)$ be the construction by Bilotta *et al.* (2012). It shows that $\omega = \omega_1\, \omega_2\, \omega_3 \ldots \omega_{2m+1}$ is the element of $CBFS_2(2m+1)$. Next, it defines $0_\omega = \{i \in \{0,1,2\} : \omega_i = 0\}$. It is a set of all positions on ω which is symbolized by 0. The researcher defines the set $CBFS_3(2m+2)$ as:

$$CBFS_3(2m+1) = \cup_{\omega \in CBFS_2(2m+1)} C^{2m+1}_{\omega,3} \qquad (4)$$

It shows that

$$C^{2m+1}_{\omega,3} = \{c \in H_3(2m+1) : 2|c_i ; \forall_{i \in 0_\omega, c_i \in \{0,1,2\}}\}$$

as a set of ternary sequences which i position is symbolized by even on {0,1,2} if the symbol of the position is 0 in ω.

From the **Construction II.C.1.1,** the researcher gets **Theorem II.C.1.2.** Code of $CBFS_3(2m+1)$ is a cross bifix free set with cardinality $2^{m+1}C_m$.

There is also the construction of ternary cross bifix free code, $CBFS_3(2m+2)$. The construction of cross bifix free set is by Affaf and Ulum (2017b).

Construction II.D.1.1. Let $CBFS_2(2m+2)$ be cross bifix free codes on Bilotta *et al.* (2012). The extension of $CBFS_2(2m+2)$ to $CBFS_3(2m+2)$ is given by putting all elements of $CBFS_2(2m+2)$ into $CBFS_3(2m+2)$ and all elements of $H_3(2m+2)$ which can be obtained from $CBFS_2(2m+2)$ by replacing 0 by 2 into $CBFS_3(2m+2)$.

From the **Construction II.D.1.1,** the researcher gets **Theorem II.D.1.2.** Set of $CBFS_3(2m+2)$ is a cross bifix free code with cardinality

$$2^m \sum_{i=0}^{\frac{m}{2}} C_m C_{m-i}$$

for m even and

$$2^m \left(\sum_{i=0}^{\frac{m}{2}} C_i C_{m-i} - C^2_{\frac{m-1}{2}} \right)$$

for m odd.

Then, it is the construction of q-ary cross bifix free code, $S^k_{q,n}$. The construction of a cross bifix free codes by Chee *et al.* (2013) is as follows:

Construction II.E.1.1. It is given a natural number n and some natural number k with $2 \leq k \leq n-2$. It denotes $S^k_{q,n}$ as set of all sequence of $s_1\, s_2\, s_3 \ldots s_n$ in $\{0,1,\cdots q-1\}^n$ which satisfies two conditions: (1) $s_1 = s_2 = s_3 = \cdots = s_k = 0$, $s_{(k+1)} \neq 0$ and $s_n \neq 0$, and (2) subsequence of $s_{(k+2)}\, s_{(k+3)}\, s_{(k+4)} \cdots s_{(n-1)}$ does not contain any string of consecutive 0.

For example, for $q = 2$, $n = 7$, and $k = 3$, the researcher obtains $S^3_{2,7}$ as the binary sequence with a length of seven. It is easy to check that binary sequence set of $S^3_{2,7}$ is {0001 111,0001011,0001101,0001001} or cross bifix free code. It should be noted that the prefix of the element in $S^k_{q,n}$ starts with consecutive zeroes and the suffix contains at most k-1 consecutive zeroes. Thus, no prefix of any length of any element can matchany suffix of itself or any other element in $S^k_{q,n}$. Therefore, $S^k_{q,n}$ must be a cross bifix free code. Next, it should consider all the possible configurations of elements in $H_q(n)$ that can be appended to the set of $S^k_{q,n}$. The researcher cannot append any element starting with a non-zero element since the non-zero element occurs in the last position of some element in $S^k_{q,n}$. Similarly, the researcher cannot append any element ending with a zero element.

There are other possible configurations of elements that the researcher needs to consider. First, let s be an element containing at least consecutive zeroes in the last n-1 position. The researcher considers the suffix starts with the last set of consecutive zeroes and contains the most k-1 consecutive zeroes following it. The suffix has the form of $0^k \alpha u$, that α is nonzero and u is a vector of length m that has the most consecutive k-1 zeroes. Then, the element of length, n, that is $0^k \alpha u 1^{n-m-k-1}$, is an element in $S^k_{q,n}$ and has a prefix matching a suffix of s. Thus, s cannot be appended to $S^k_{q,n}$.

Second, it lets s be an element which contains a prefix at most of k-1 zeroes followed by a non-zero element, that is $s = 0^l \alpha u$. It shows that α is non-zero, $0 < l \leq k-1$, and u has the length of n-l-1. It is readily seen that $0^l \alpha$ is also the suffix of the element in $0^k 1^{n-k-l-1} 0^l \alpha$ in $S^k_{q,n}$. Hence, the element in $H_q(n)$ cannot be appended to $S^k_{q,n}$.Thus, no additional element can be appended to the set $S^k_{q,n}$, while it still preserves the cross-bifix-free property.

From the **Construction II.E.1.1,** the researcher obtains **Theorem II.E.1.2.** It gives the natural number n and some natural number k with $2 \leq k \leq n-2$. Set of $S^k_{q,n}$ is a cross bifix free code with cardinality $(q-1)^2 F_{k,q}(n-k-2)$ which is non-expandable on $H_q(n)$, and $F_{k,q}(n)$ sequence satisfies $F_{k,q}(n) = (q-1) \sum_{l=1}^{k} F_{k,q}(n-l)$ with the conditional value of $F_{k,q}(0), F_{k,q}(1), F_{k,q}(2), \ldots, F_{k,q}(k-1)$.

Moreover, there are upper bound and optimal codes. First, the upper bound of cross bifix free code will be explained. Chee *et al.* (2013) revisited the construction in Bajic (2007). They gave a new construction of cross-bifix-free code that generalizes the construction in two ways. Firstly, they provided new binary codes that were greater in cardinality compared to the ones in Bilotta *et al.* (2012) for larger lengths. In the process, they discovered the interesting connections of the size of the codes obtained in the so-called k-generalized Fibonacci number. Secondly, they generalized the construction of q-ary alphabets for any $q \geq 2$. To the best of their knowledge, this was the first construction of cross bifix free codes over alphabets of size greater than two. The size of the generalized q-ary constructions was also related to a Fibonacci sequence, which they called the (q-1)-weighted k-generalized Fibonacci sequence. Using this relation to the Fibonacci sequences, Chee *et al.* (2013) analyzed the asymptotic size of their construction. In the process of this asymptotic analysis, they generalized a result of Dresden and Du (2014) on k-generalized Fibonacci sequence to (q-1)-weighted k-generalized Fibonacci sequence. They let $C(n,q)$ denote the optimal size of a cross bifixfree code of length n over an alphabet of size q.

An upper bound for the optimal size of a cross bifix free code is readily obtained from the research of the statistical properties of the sets in the data stream. The main object of research is the time when searching for any word of the cross bifix free code in the data stream returns with a positive match. From the information, Chee *et al.* (2013) got the result that the upper bound of optimal code size was no more than $\frac{q^n}{2n-1}$, that is

$$\mathbb{C}(n,q) \leq \frac{q^n}{2n-1} \tag{5}$$

However, Blackburn (2015) stated that the optimal code size would never reach A. In **Theorem II.F.1.1,** it lets n and q be integers with $n \geq 2$ and $q \geq 2$. It also lets $C(n,q)$ be the number of codewords in the largest cross bifix free codes of length n and symbolic q, it becomes as follows:

$$\mathbb{C}(n,q) < \frac{q^n}{2n-1} \tag{6}$$

In obtaining these results, Blackburn (2015) looks at the set of C_q^i that is:

$$C_q^i = \{(\omega,i) : \omega = \omega_1 \dots \omega_{2n-1} \in F^{2n-1},$$

$$\omega_i \dots \omega_{i+n-1} \in C(n,q),\ i \in [2n-1]\} \tag{7}$$

It shows that $|F| = q$, $C(n,q)$ is a cross bifix free code with optimal size, and $[2n-1]$ is $\{1,2,\dots,2n-1\}$. For example, (11101010010110,2) is an element of C_2^2. It should consider that two elements in C_q^i cannot appear as distinct cyclic subsequence of any ω of length $2n$-1. Thus, for any ω in F^{2n-1}, there is at most one choice for an integer i such that (ω,i) in C_q^i. It is clear that it is $|C_q^i| = (2n-1)\ C(n,q)q^{n-1}$ since there are $2n$-1 choice for i, $C(n,q)$ choices for the codeword starting in the i^{th} position of ω, and q^{n-1} choices for the remaining positions in ω. Moreover, no subsequence of any of the qconstant words ω of length in $2n$-1 can appear as a codeword in $C(n,q)$. So, the researcher gets $|C_q^i| \leq q^{2n-1} - q < q^{2n-1}$. Finally, from $|C_q^i| = (2n-1)\ C(n,q)q^{n-1}$, it will be concluded that $(2n-1)\ C(n,q)\ q^{n-1} < \frac{q^n}{2n-1}$, is $C(n,q) < \frac{q^n}{2n-1}$.

Second, it is optimal size of cross bifix free code. Blackburn (2015) gave the construction of cross bifix free codes C by using the construction of Chee *et al*. (2013). Blackburn defines it as **Definition II.F.2.1.** It gives the natural number k and non-empty set F with cardinality q. In $S \subseteq F^k$, $c_1\ c_2\ c_3 \dots c_r \in F^r$ is S-free if and only if $r < k$ or if $r \geq k$, and $c_i\ c_{i+1}\ c_{i+2} \dots c_{i+k-1} \notin S$ is for every i in $\{1, 2, \dots, r\text{-}k+1\}$. By using **Definition II.F.2.1,** Blackburn (2015) modified the construction of Chee *et al*. (2013) by Construction II.F.2.2. It gives a natural number l and some natural number k with $1 \leq k \leq n$-1 and $1 \leq l \leq q$-1. It lets $F = I \cup J$ be a partition of a set Fof cardinality q into two parts I and J of cardinalities l and q-l. If $S \subseteq I^k \subseteq F^k$ and C are the set of all $c = c_1\ c_2\ c_3 \dots c_n \in F^k$ like (1) $c_1\ c_2\ c_3 \dots c_K \in S$, (2) $c_{k+1} \in S$ and $c_n \in S$, and (3) $c_{k+2}\ c_{k+3} \dots c_{n-1}$ is S-free, there is cross bifix free code.

From the **Construction II.F.2.2,** Blackburn (2015) got **Theorem II.F.2.3.** It lets n and q be positive integers like $n \geq 2$ and q ≥ 2. It lets the largest cross bifix free code have cardinality $C(n,q)$. When n divides q, the researcher has:

$$\mathbb{C}(n,q) = \frac{1}{n}\left(\frac{n-1}{n}\right)^n q^n \tag{8}$$

III. RESULTS AND DISCUSSIONS

In maximality on construction $CBFS_3(n)$, Affaf (2018) explained that $|CBFS_3(3)|$ was close to the optimal cardinality that it was possible to achieve maximum cardinality. On the other side, in the construction by Affaf and Ulum (2017a), the researcher swaps symbol 1 with 0 on $CBFS_2$ (2m+1) and extends it to $CBFS_3(2m+1)$. However, it does not pass the expansion of $CBFS_2(2m+2)$ to $CBFS_3(2m+2)$. In this research, it is assumed that the researcher swaps symbol 1 with 0 on $CBFS_3(2m+1)$ and $CBFS_3(2m+2)$. For example, for 1111000 on $CBFS_3(2m+2)$ and 111000 on $CBFS_3(2m+2)$, the researcher writes those as 0000111 and 000111 respectively.

In Affaf and Ulum (2017b), $CBFS_3(2m+2)$ on **Construction II.C.1.1.** is also shown as

$$CBFS_3(2m+2) = \cup_{\omega \in CBFS_2(2m+2)} C_{\omega,3}^{2m+2} \tag{9}$$

In other word, the researcher can simplify $CBFS_3(2m+1)$ and $CBFS_3(2m+2)$ as $CBFS_3(n)$, that is

$$CBFS_3(n) = \cup_{\omega \in CBFS_2(n)} C_{\omega,3}^{n} \tag{10}$$

$$C_{\omega,q}^{n} = \{c \in H_3(n) : 2|c_i, \forall_{c_i \in \{0,1,2\}, i \in 0_\omega}\} \tag{11}$$

Furthermore, the research will claim that $CBFS_3(n)$ is non-expandable. This result is stated on **Theorem III.A.1.1.** For $n \geq 1$, $CBFS_2(n)$ is cross bifix free code which is non-expandable on $H_3(n)$.

For the proof, Since all elements of $CBFS_2(n)$ are started by 0 and finished by 1 and $C_{\omega,3}^{n}$ set of all ternary sequences, the i^{th} position is an even on $\{0,1,2\}$. If the position is 0 on ω and the i^{th} position is odd on $\{0,1,2\}$, the position is 1 on ω for all $\omega \in CBFS_2(n)$, and all elements h in $H_3(n)$ are started bv svmbolic odd or finished by symbolic even, there is z in $C_{\omega,3}^{n}$ such that $pre_1\ h = suf_1 z$ or $suf_1 h = pre_1 z$, respectively. Thus, it is enough to show that $CBFS_3(n)$ is non-expandable on $\hat{H}_3(n)$. It is all ternary sequence with length n which is started by even symbol and finished by an odd symbol.

Then, it is $h\ \hat{H}_3(n)$. The $\bar{h}$ is binary sequence with length n obtained from h by replacing all even symbol on h by 0 and all odd symbol by 1. For sure, $\bar{h}$ is an element of $H_2(n)$. Furthermore, by **Theorem II.A.2.1, Theorem II.B.2.1,** and **Theorem II.C.2.1,** $CBFS_2(n)$ is non-expandable on $H_2(n)$. In other words, there is ω in $CBFS_2(n)$ like $pre_k\ \omega = suf_k\ \bar{h}$ or $suf_k \omega = pre_k\ \bar{h}$ for some natural number k which satisfies $1 < k < n$. Therefore, there is c $\in C_{\omega,3}^{n}$ like pre_k c $= suf_k h$ or $suf_k c = pre_k h$. So, $CBFS_3(n)$ is cross bifix free code which is non-expandable on $H_3(n)$.

Next, Theorem III.A.I.I states that $CBFS_3(n)$ is maximal for arbitrary length. In this part, it will be shown that for $n = 3$, $CBFS_3(n)$ is optimal. This result states **Theorem III.B.I.I.** For $n = 3$, $CBFS_3(n)$ is optimal. From Blackburn (2015), the researcher knows that $C(3,q)$, the number of optimality of cross bifix free code with q symbol and length $n = 3$, is equal to $\left[\frac{2q}{3}\right]^2\left(q - \left[\frac{2q}{3}\right]\right)$. So, for $q = 3$, the researcher gets $\mathbb{C}(3,q) = \left[\frac{6}{3}\right]^2\left(3 - \left[\frac{6}{3}\right]\right)$. On the other side, the cardinality of $CBFS_3(3)$ is equal to $2^{1+1}C_1$ according to reference Affaf and Ulum (2017a). Finally, the researcher concludes that $C(3,q)=|CBFS_3(3)|$ is for $q = 3$. So, $CBFS_3(3)$ is optimal.

In the comparison cardinality of $CBFS_3(n)$ and the optimal code, $C(n,3)$shows that the optimal cardinality of the ternary cross bifix free codes with length n, and $|CBFS_3(n)|$ mentions that the cardinality of the ternary cross bifix free code $CBFS_3(n)$. Using Stirling's approximation, the researcher obtains that the number C_m is approximate:

$$C_m \approx \frac{1}{m+1}\frac{2^{2^m}}{\sqrt{\pi m}} \tag{12}$$

So, the researcher gets:

$$\frac{\mathbb{C}(n,3)}{|CBFS_3(n)|} \approx \frac{\frac{3^n}{2n-1}}{2^n \frac{2^{2^m}}{(m+1)\sqrt{m\pi}}} \tag{13}$$

That is,

$$\frac{\mathbb{C}(n,3)}{|CBFS_3(n)|} \approx \frac{2(m+1)\sqrt{m\pi}}{2m+1} \tag{14}$$

From here, it can be seen that $|CBFS_3(n)|$ is near-optimal cardinality so that $|CBFS_q(n)|$ for arbitrary q is very likely to achieve the maximum.

IV. CONCLUSIONS

In this research, the researcher has shown that the construction of $CBFS_3(n)$ on Affaf and Ulum (2017a, 2017b) achieves the maximality, and it is non-expandable in $H_3(n)$. Furthermore, for $n = 3$, $CBFS_3(n)$, it reaches the optimality. It can be seen that $|CBFS_3(n)|$ is near-optimal cardinality so that $|CBFS_3(q)|$ for arbitrary q is very likely to achieve the maximum too. It means that future research can be done by exploring whether for any q or the code $CBFS_3(q)$ will also be maximal.

ACKNOWLEDGEMENT

This research was supported by a grant from STKIP PGRI Bangkalan in 2018. The author felt indebted to the LPPM STKIP PGRI Bangkalan which provided a grant to assist this research.

REFERENCES

Affaf, M. (2018). Kemaksimalan kode cross bifix bebas. In *Seminar Nasional Pendidikan Matematika Ahmad Dahlan* (Vol. 6).

Affaf, M., & Ulum, Z. (2017a). Konstruksi kode cross bifix bebas ternair untuk panjang ganjil. In *Prosiding SI MaNIs (Seminar Nasional Integrasi Matematika dan Nilai-Nilai Islami)* (Vol. 1, No. 1, pp. 1-5).

Affaf, M., & Ulum, Z. (2017b). Konstruksi kode cross bifix bebas ternari berpanjang genap untuk mengatasi masalah sinkronisasi frame. *JIKO (Jurnal Informatika dan Komputer), 2*(2), 109-116.

Bajic, D. (2007). On construction of cross-bifix-free kernel sets. In *2nd COST2100 MCM*.

Bajic, D., & Loncar-Turukalo, T. (2014). A simple suboptimal construction of cross-bifix-free codes. *Cryptography and Communications, 6*(1), 27-37.

Bernini, A., Bilotta, S., Pinzani, R., Sabri, A., & Vajnovszki, V. (2014). Prefix partitioned gray codes for particular cross-bifix-free sets. *Cryptography and Communications, 6*(4), 359-369.

Bilotta, S., Grazzini, E., Pergola, E., & Pinzani, R. (2013). Avoiding cross-bifix-free binary words. *Acta Informatica, 50*(3), 157-173.

Bilotta, S., Pergola, E., & Pinzani, R. (2012). A new approach to cross-bifix-free sets. *IEEE Transactions on Information Theory, 58*(6), 4058-4063.

Blackburn, S. R. (2015). Non-overlapping codes. *IEEE Transactions on Information Theory, 61*(9), 4890-4894.

Chee, Y. M., Kiah, H. M., Purkayastha, P., & Wang, C. (2013). Cross-bifix-free codes within a constant factor of optimality. *IEEE Transactions on Information Theory, 59*(7), 4668-4674.

Dresden, G. P., & Du, Z. (2014). A simplified Binet formula for k-generalized Fibonacci numbers. *Journal of Integer Sequences, 17*(4), 1-9.

Levy, M., & Yaakobi, E. (2017). Mutually uncorrelated codes for DNA storage. In *2017 IEEE International Symposium on Information Theory (ISIT)*(pp. 3115-3119).

Massey, J. (1972). Optimum frame synchronization. *IEEE Transactions on Communications, 20*(2), 115-119.

Nielsen, P. (1973). On the expected duration of a search for a fixed pattern in random data. *IEEE Transactions on Information Theory, 19*(5), 702-704.

Van Wijngaarden, A. D. L., & Willink, T. J. (2000). Frame synchronization using distributed sequences. *IEEE Transactions on Communications, 48*(12), 2127-2138.

Weindl, J., & Hagenauer, J. (2007). Applying techniques from frame synchronization for biological sequence analysis. In *2007 IEEE International Conference on Communications* (pp. 833-838). IEEE.