2009

ANALYSIS OF SECURITY MEASURES FOR SEQUENCES

Ramakanth Kavuluru
University of Kentucky, kvnramakanth@yahoo.com

ABSTRACT OF DISSERTATION

Ramakanth Kavuluru

The Graduate School
University of Kentucky
2009

ANALYSIS OF SECURITY MEASURES FOR SEQUENCES

ABSTRACT OF DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Engineering at the
University of Kentucky

By
Ramakanth Kavuluru
Lexington, Kentucky

Director: Dr. Andrew Klapper, Professor of Computer Science
Lexington, Kentucky 2009

ABSTRACT OF DISSERTATION

ANALYSIS OF SECURITY MEASURES FOR SEQUENCES

Stream ciphers are private key cryptosystems used for security in communication and data transmission systems. Because they are used to encrypt streams of data, it is necessary for stream ciphers to use primitives that are easy to implement and fast to operate. LFSRs and the recently invented FCSRs are two such primitives, which give rise to certain security measures for the cryptographic strength of sequences, which we refer to as complexity measures henceforth following the convention. The linear (resp. N-adic) complexity of a sequence is the length of the shortest LFSR (resp. FCSR) that can generate the sequence. Due to the availability of shift register synthesis algorithms, sequences used for cryptographic purposes should have high values for these complexity measures. It is also essential that the complexity of these sequences does not decrease when a few symbols are changed. The k-error complexity of a sequence is the smallest value of the complexity of a sequence obtained by altering k or fewer symbols in the given sequence. For a sequence to be considered cryptographically 'strong' it should have both high complexity and high error complexity values.

An important problem regarding sequence complexity measures is to determine good bounds on a specific complexity measure for a given sequence. In this thesis we derive new nontrivial lower bounds on the k-operation complexity of periodic sequences in both the linear and N-adic cases. Here the operations considered are combinations of insertions, deletions, and substitutions. We show that our bounds are tight and also derive several auxiliary results based on them.

A second problem on sequence complexity measures useful in the design and analysis of stream ciphers is to determine the number of sequences with a given fixed (error) complexity value. In this thesis we address this problem for the k-error linear complexity of 2^n-periodic binary sequences. More specifically:

1. We characterize 2^n-periodic binary sequences with fixed 2- or 3-error linear complexity and obtain the counting function for the number of such sequences with fixed k-error linear complexity for $k = 2$ or 3.

2. We obtain partial results on the number of 2^n-periodic binary sequences with fixed k-error linear complexity when k is the minimum number of changes required to lower the linear complexity.

Keywords: stream ciphers, LFSRs, FCSRs, sequence complexity measures, k-error complexity

Author's signature: Ramakanth Kavuluru

Date: September 3, 2009

ANALYSIS OF SECURITY MEASURES FOR SEQUENCES

By
Ramakanth Kavuluru

Director of Dissertation: Andrew Klapper

Director of Graduate Studies: Raphael Finkel

Date: September 3, 2009

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name	Date

DISSERTATION

Ramakanth Kavuluru

The Graduate School
University of Kentucky
2009

ANALYSIS OF SECURITY MEASURES FOR SEQUENCES

DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Engineering at the
University of Kentucky

By
Ramakanth Kavuluru
Lexington, Kentucky

Director: Dr. Andrew Klapper, Professor of Computer Science
Lexington, Kentucky 2009

To my mother, Lakshmi Kavuluru.
Thank you amma, for all the love you shower on me.

ACKNOWLEDGMENTS

I am fortunate to have had Dr. Andrew Klapper as the advisor for my doctoral studies. I started in the program without having some essential mathematical background. Dr. Klapper has patiently taught me the basics of cryptography and sequences. His excellent guidance in technical writing will be useful for the rest of my career. This thesis would not have been possible without his prompt and authoritative responses to my numerous questions on the subject. I thank him for inspiring and guiding me throughout the program.

My special thanks to Dr. Judy Goldsmith for providing sensible advice during my job search. Her detailed comments on an initial draft of the thesis have greatly helped improve the presentation of the material. I would also like to thank Dr. Mirosław Truszczyński and Dr. David Leep for serving on my committee and for providing great career advice whenever I approached them. Thanks also to Dr. Margaret Readdy for acting as the external examiner for the final exam and some important comments on the thesis. My sincere thanks to Dr. Edgar Enochs of the mathematics department at the University of Kentucky for teaching me abstract algebra and clarifying all my questions on finite fields. I also thank Erik Stokes of Michigan Technological University for providing the LaTeX framework used in the preparation of this document.

My first inspiration to pursue research in computer science came from my interactions with Dr. Uta Ziegler. She always gave her best and challenged students in the courses I took with her during my masters program at the Western Kentucky University. The commitment she showed when I was working with her as a research assistant inspired me to go further in computer science. I am grateful to her and the faculty who taught me at the Western Kentucky University for stimulating further interest in the area.

There are several other people in the computer science department at the University of Kentucky who were very supportive. I would like to thank Dr. Kenneth Calvert, Dr. Zongming Fei, and Dr. Jim Griffioen for kindly providing office space and resources. Many thanks to Dr. Greg Wasilkowski for providing excellent support during the job searching phase. Thanks to Santosh Chandrasekhar for being very helpful by spending his time and energy to give me rides to near-by presentations and interviews. I am also thankful for the time and efforts of all the staff members in the department.

I am very grateful to my parents, sisters, and friends for their support and encouragement. I thank my father, Satyanarayana Kavuluru, for heartily encouraging me to pursue a Ph.D and having faith in my abilities, even before I started in the program. My mother, Lakshmi Kavuluru, has been a great example of the values I embrace and I thank her profusely for instilling them in me. My sisters, Shalini and Manjusha, always cheered me up and helped me relax whenever I talked with them.

I had wonderful roommates during my masters program who have taken great interest in my progress and stood by me through thick and thin. My special thanks to Pradeep Arumalla, Srikanth Bodla, Naveen Boppana, and Sridhar Maguluri for being much more than roommates. Thanks to Chaitu, Pavan, Sharath, Vissu, and all the members of Lexington Sai center for all the good times and great support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Introduction

Over the past three decades digital computers and computer networks have revolutionized the ways in which information is processed and communicated. Wireless and sensor networks have further improved our ability and made it more convenient to acquire and process data and communicate information. The digital revolution has had a significant impact on all spheres of human life. Digital computing and communication have become indispensable for making rapid progress in physical and biological sciences, engineering, and even in arts and sports.

While computers and other electronic marvels mostly hastened the advancement of science and technology, the past fifteen years has marked a sharp increase in the way they are used by individuals once they became affordable for personal use. The advent of PCs, the Internet, and cellular phones has made it easy for people to communicate with others and do activities like paying bills, buying merchandise, and handling financial transactions. While this has tremendously increased convenience and productivity, it has also created a great risk due to accidental disclosure of, or malicious attempts to gain access to, sensitive information. Instead of a few watchful and needfully paranoid security experts, the current situation requires all individuals to be careful when using digital communication.

In this chapter we introduce concepts and ideas fundamental to secure communication, which form the basis for this thesis.

1.1 Cryptography and Stream Ciphers

Cryptography is the mathematical study of techniques and tools to hide information and to communicate over insecure channels so that it is infeasible for an eavesdropper to understand what is being communicated. While cryptography has been used since at least ancient Roman times for military purposes, in the past three decades it has become essential for the functioning of the modern digital society.

Suppose a sender wants to send a message to a receiver securely. In all cryptosystems the sender and receiver have some secret information — so called encryption and decryption keys. The sender uses her key to scramble the message in such a way that (hopefully) only the legitimate receiver who has the valid decryption key can unscramble it. A private key cryptosystem is a cryptographic scheme where encryption and decryption keys are identical. Stream ciphers are private key cryptosystems used for security in settings where very high speed is essential and the users can accept a suboptimal level of security. They are typically used in digital telephones, video on demand, and other applications where the volume of data being transmitted is very high. In a stream cipher the message is treated as a sequence of symbols from a fixed alphabet, usually either the binary alphabet $\{0, 1\}$ or the set of all bytes or words. The key stream is also a sequence of symbols, usually from the same alphabet used for message streams. The sender encrypts each message symbol by combining it with the next key stream symbol using modular arithmetic. The resulting new sequence,

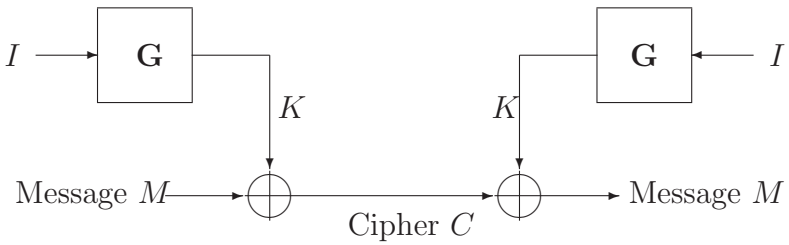

Figure 1.1: Stream Cipher Schematic

called the cipher, is sent across the channel to the receiver. The receiver uses the same key stream to decrypt the cipher by performing the opposite operation to that performed by the sender to get back message symbols. All of the security of a stream cipher comes from the design and analysis of the generator of the key stream.

Although messages and key streams can be treated as composed of 1s and 0s, the actual generation of the key stream can usually be over symbols from a finite field $\mathbb{F}_q$ of size q where q is a power of a prime. For simplicity, in this section we restrict our attention to the binary field $\mathbb{F}_2 = \{0, 1\}$ with addition and multiplication modulo 2. The schematic for a stream cipher is shown in Figure 1.1. A generator G with an initialization seed I (which acts as the key for the generator) is used to generated the key stream K for encryption. Since the key stream used for both encryption and decryption should be same, an identical generator is used at the receiver's end. Let $M = m_0, m_1, \ldots$ be the message symbol sequence and let $K = k_0, k_1, \ldots$ be the key stream. The sender forms and transmits the cipher $C = c_0, c_1, \ldots$ where $c_i = m_i \oplus k_i$, for $i = 0, 1, \ldots$. The symbol $\oplus$ denotes addition modulo 2 (or XOR) given by $0 \oplus 1 = 1 \oplus 0 = 1$ and $1 \oplus 1 = 0 \oplus 0 = 0$. The receiver recovers the message by performing another XOR operation between the same key stream K and the cipher C to obtain $m_i = c_i \oplus k_i$, where $i = 0, 1, \ldots$. An attacker who has access to the cipher that is transmitted cannot, the sender hopes, understand the message if he does not know the key stream. The typical assumption in the analysis of the security of stream ciphers is that the attacker has access to a part of the key stream (this can be found by knowing a piece of the message and the corresponding piece of the cipher) and wants to use this to predict the remainder of the key stream. Thus, the problem of designing a good stream cipher is reduced to the problem of designing a fast key stream generator whose full output is hard to predict from a prefix of the output.

1.2 Thesis Organization

Linear feedback shift registers (LSFRs)—see Section 2.1—are widely used as components in key stream generators for use in stream ciphers. Recently, FCSRs are also being used to build stream ciphers [3]. In Chapter 2 we discuss the architecture of LFSRs and FCSRs and present some properties that make them suitable for use in stream ciphers. We also introduce linear complexity and N-adic complexity, the

sequence security measures that arise from LFSRs and FCSRs respectively. We give formal definitions and also discuss several variants of these measures providing the motivation for each measure. We list the standard problems that are considered for sequence complexity measures and present some classical results on this topic.

Chapters 3 and 4 form the main contribution of this thesis[1]. In Chapter 3 we present our results on lower bounds for the error linear complexity and the error N-adic complexity of periodic sequences. We show that there exist infinite families of sequences that achieve the bounds in the linear case. These results also give new nontrivial lower bounds on the minimum number of operations required to lower the complexity for both the linear and N-adic cases.

In Chapter 4 we obtain counting functions for the k-error linear complexity of 2^n-periodic binary sequences. We analyze the Games-Chan algorithm and obtain some properties of the structure of 2^n-periodic binary sequences. For a given L and an integer $\mathcal{C}$ in a certain range, we determine the counting function for the number of such sequences with linear complexity L and k-error linear complexity $\mathcal{C}$ when k is the minimum number of changes needed to lower the linear complexity of sequences with linear complexity L. We also determine the counting function for the number of 2^n-periodic binary sequences with fixed 2-error or 3-error linear complexity. Throughout the chapter we motivate the problems we solve and survey results on complexity measures for prime power periodic sequences.

In Chapter 5 we discuss future research directions on sequence complexity measures. We give a brief account of stream cipher design and analysis and list problems we intend to pursue in the future.

[1]This thesis is based upon work supported, in part, by the National Science Foundation under Grant No. CCF-0514660. Any opinions, findings, and conclusions or recommendations expressed in this thesis are those of the author and do not necessarily reflect the views of the National Science Foundation

2 Background and Preliminaries

In this chapter we introduce LFSRs, FCSRs, and some of their properties. For detailed expositions on LFSRs and their properties please refer to the books by Golomb [20], Golomb and Gong [21], and McEliece [50]. An upcoming book by Goresky and Klapper [23] covers both LFSRs and FCSRs. Based on these shift registers we introduce sequence complexity measures, which form the basis for the main results in the thesis. We discuss the motivation for studying these complexity measures and present some well known results on them.

2.1 LFSR and FCSR Basics

LFSRs have been used for at least 50 years as building blocks for a wide variety of communications and computing applications, including stream ciphers, error correcting codes, CDMA spread spectrum communication, and quasi-Monte Carlo applications. LFSRs are fast and simple to implement in hardware. The statistical properties of LFSR sequences are also thoroughly studied using well known algebraic methods, especially the theory of finite fields.

LFSRs are used to generate sequences that satisfy homogeneous linear recurrence relations over finite fields. (Please see the book by Lidl and Niederreiter [45] for details of finite field theory.) Let $\mathbb{F}_q$ be the finite field with $q = p^r$ elements where p is a prime and r is a positive integer. An LFSR has a fixed number of cells each loaded with an element in $\mathbb{F}_q$ and tapped by using an element of $\mathbb{F}_q$. Let m be the number of cells in an LFSR and let $c_1, \ldots, c_m \in \mathbb{F}_q$ be their taps as shown in Figure 2.1. In each step the LFSR operates by shifting the contents of the register to the right by one cell and the right most element s_{n-m} is output as the next element of the sequence. The new element s_n fed back into the left most cell is computed by

$$s_n = \sum_{i=1}^{m} c_i s_{n-i},$$

where the summation is using addition in $\mathbb{F}_q$. The state of the LFSR at any particular step is given by the tuple $(s_{n-m}, \ldots, s_{n-1})$.

Definition 2.1. A sequence $\mathbf{S} = (s_0, s_1, \ldots)$ is called eventually periodic if and only if there exist integers $r > 0$ and $k \geq 0$ such that $s_{n+r} = s_n$ for all $n \geq k$. The smallest such r is called the least period of $\mathbf{S}$. If there exists such an r with $k = 0$, then $\mathbf{S}$ is called strictly periodic or just periodic.

Since the number of possible states for an LFSR of a fixed length m is q^m, sequences produced by LFSRs are eventually periodic as a state must repeat after q^m states. Conversely any eventually periodic sequence can be generated by some LFSR. We can also see that a sequence generated by an LFSR is periodic if the tap on the right most cell is nonzero. The maximum period of any sequence generated by an

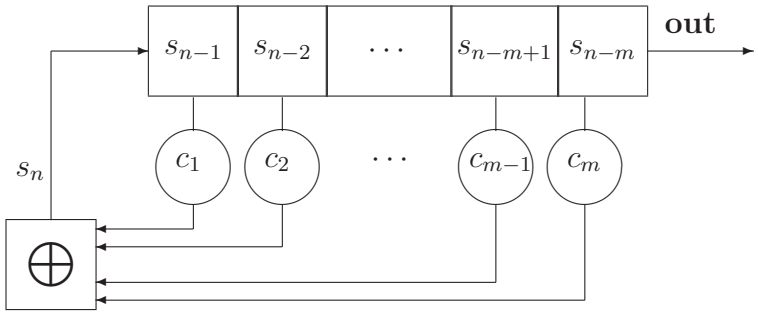

Figure 2.1: A Linear Feedback Shift Register of Length m.

LFSR of length m is $q^m - 1$: the all zero state cannot be included since it would result in a sequence of period 1.

We associate a connection polynomial $f(x)$ to an LFSR based on its taps as

$$c(x) = c_0 + c_1 x + \ldots + c_{m-1}x^m, \text{ where } c_0 = -1,$$

which is useful to analyze sequences generated by the LFSR. We enumerate some of the properties of LFSRs. A polynomial $c(x)$ of degree m in $\mathbb{F}_q$ is called primitive if it has a root in $\mathbb{F}_{q^m}$ that is a primitive element of $\mathbb{F}^*_{q^m} = \mathbb{F}_{q^m} \setminus \{0\}$.

(i) Every period of any sequence generated by the LFSR with connection polynomial $c(x)$ divides every $T \geq 1$ such that $c(x)|x^T - 1$.

(ii) The power series $\sum_{i\geq 0} s_i x^i$ associated with a sequence $\mathbf{S} = (s_0, s_1, \ldots)$ generated by an LFSR is a rational function over $\mathbb{F}_q[x]$ of the form $g(x)/c(x)$. The sequence is periodic if and only if $\deg(g(x)) < \deg(c(x))$.

(iii) An LFSR sequence with maximal period $2^m - 1$ is called an m-sequence. M-sequences are sequences generated by LFSRs whose connection polynomials are primitive.

While LFSRs are simple and efficient, researchers have also been looking for other efficient ways of generating pseudorandom sequences. One such way is to add a small amount of memory to the basic shift register architecture that can be used as a "carry" in the calculations. This idea is originally motivated by the summation combiner [75], which adds two binary sequences using addition with carry (as opposed to addition modulo 2) in an attempt to produce hard-to-predict sequences for cryptographic purposes. To analyze the summation combiner Klapper and Goresky [41] introduced the idea of adding memory into the usual LFSR structure and thus invented FCSRs in 1993. It should be noted here that the same idea was introduced for random number generation around the same time by Couture and L'Ecuyer [10] and Marsaglia [47].

FCSRs are thus an arithmetic or with-carry analog of LFSRs. An FCSR generates sequences over $\{0, \ldots, N-1\}$ for some $N \geq 2$. The analysis of FCSRs is based on algebra over the N-adic numbers.

Definition 2.2. An N-adic number is an infinite expression

$$a = \sum_{i=0}^{\infty} a_i N^i,$$

where $a_0, a_1, \ldots \in \{0, \ldots, N-1\}$.

The set of N-adic numbers is denoted by $\mathbb{Z}_N$. The addition and multiplication are defined to take carries into account, which is the main difference in the corresponding operations over the ring of formal power series in x. With these operations $\mathbb{Z}_N$ is a ring and the additive inverse of the multiplicative identity element is $-1 = (N-1) + (N-1)N + (N-1)N^2 + \cdots$.

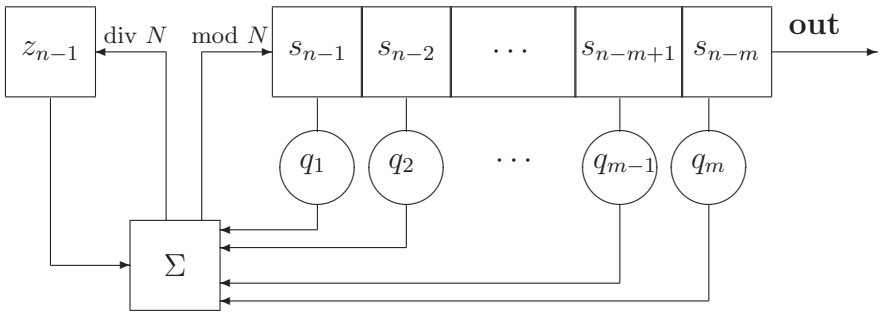

Figure 2.2: A Feedback With Carry Shift Register of Length m.

The architecture of an FCSR is similar to that of an LFSR with the exception of a memory cell that is used in the computation of the element fed back in each step. Figure 2.2 shows the architecture of an FCSR with m cells. The state of the FCSR is determined by the contents of the m cells and the value in the memory register and is represented by $(s_{n-m}, \ldots, s_{n-1}; z_{n-1})$ where $s_{n-i} \in \{0, \ldots, N-1\}$ and $z_{n-1} \in \mathbb{Z}$.

In each step the state change operation is described as follows.

(i) Compute the integer sum

$$\sigma_n = \sum_{i=1}^{m} q_i s_{n-i} + z_{n-1}.$$

(ii) Shift the contents one step to the right and output the element in the right most cell.

(iii) Put $s_n = \sigma_n \mod N$ in the left most cell.

(iv) Put $z_n = \lfloor \sigma_n / N \rfloor$ in the memory cell.

We associate the connection number

$$q = q_0 + q_1 N + \cdots + q_m N^m, \text{ where } q_0 = -1,$$

corresponding to the m taps. We have the following properties for FCSRs.

(i) The N-adic number $\sum_{i \geq 0} s_i N^i$ associated with a sequence $\mathbf{S} = (s_0, s_1, \ldots)$ generated by an FCSR is a rational number of the form $-p/q$ where q is the connection number of the FCSR. We have the following facts about the periodicity of $\mathbf{S}$.

   a) $\gcd(N, q) = 1$.

   b) $\mathbf{S}$ is eventually periodic.

   c) $-p/q = 1$ if and only if $\mathbf{S} = (N-1, N-1, \ldots)$

   d) $\mathbf{S}$ is periodic if and only if $0 \leq p \leq q$.

(ii) Every period of any sequence generated by the FCSR with connection number q divides every $T \geq 1$ such that $q | N^T - 1$. In particular, the least period divides the order of N modulo q.

(iii) From (ii) the maximal period of an FCSR sequence is $q - 1$ which is achieved if and only if N is primitive modulo q and q is prime. L-sequences are those sequences generated by FCSRs with prime connection numbers and with N primitive modulo the connection number.

Note that the memory component in FCSRs is an integer. So for practical purposes memory should be bounded over any infinite execution of an FCSR, which turns out to be the case as shown by Klapper and Goresky [23, Theorem 7.3.2].

2.2 Linear Complexity and N-adic Complexity Measures

LFSRs and FCSRs give rise to security measures that determine the unpredictability of a sequence based on the size of the registers that can generate the sequence. In this section we explore these measures, their properties, and several interesting results and problems about them.

2.2.1 Linear Complexity

Linear complexity and related measures have been explored extensively over the past five decades. We start with a definition that is important from an engineering point of view.

Definition 2.3. Let $\mathbf{S} = (s_0, s_1, \ldots)$ be a finite or infinite sequence over $\mathbb{F}_q$. The linear complexity $L(\mathbf{S})$ of $\mathbf{S}$ is the length of smallest LFSR that can generate $\mathbf{S}$.

We note that if $\mathbf{S}$ in Definition 2.3 is a finite string of length n, then $L(\mathbf{S})$ denotes the length of the shortest LFSR whose first n output symbols coincide with $\mathbf{S}$.

For the rest of the document let $f^*(x)$ denote the reciprocal polynomial of $f(x)$ defined as $f^*(x) = x^{\deg f} f(1/x)$. We recall that a recurrence relation satisfied by the sequence $\mathbf{S} = (s_0, s_1, \dots)$ generated by the LFSR of length m with connection polynomial $c(x) = \sum_{i=1}^{m} c_i x^i - 1$ is

$$s_n - (c_1 s_{n-1} + \cdots + c_m s_{n-m}) = 0, \text{ for all } n \geq m.$$

We call the polynomial $c^*(x)$ the characteristic polynomial associated with this recurrence relation where $c(x)$ is the connection polynomial of the LFSR. A sequence might satisfy more than one recurrence relation and hence can have many characteristic polynomials. The unique monic characteristic polynomial with the least degree is called the minimal polynomial of $\mathbf{S}$. Hence the linear complexity can also be defined as the least order of a homogeneous linear recurrence relation satisfied by the sequence. If $c_m \neq 0$, that is, for a periodic sequence, the linear complexity can also be defined as the degree of its minimal polynomial.

Let $\mathbf{S}$ be a periodic sequence of period T and let $\mathbf{S}(x) = s_0 + s_1 x + \cdots + s_{T-1} x^{T-1}$ be the polynomial corresponding to its first period $(s_0, s_1, \dots, s_{T-1})$. We can see that

$$\sum_{i=0}^{\infty} s_i x^i = \frac{\mathbf{S}(x)}{1 - x^T}.$$

If $g(x)/f(x)$ is the reduced form of $\mathbf{S}(x)/(1-x^T)$ so that $\gcd(f(x), g(x)) = 1$, we call $f(x)$ the minimal connection polynomial and $f^*(x)$ the minimal polynomial. We can also express the linear complexity as $L(\mathbf{S}) = T - \deg(\gcd(\mathbf{S}(x), 1 - x^T))$.

The smallest LFSR that generates a given sequence can be determined using the Berlekamp-Massey algorithm [48] using only the first $2L$ elements of the sequence in $\mathcal{O}(L^2)$[1] field operations, where L is the linear complexity of the sequence. Hence, for cryptographic purposes, sequences with high linear complexity are essential, as an adversary would then need large initial segments of the sequences to recover the LFSRs that generate them using the Berlekamp-Massey algorithm. It is important to note that the algorithm is adaptive in the sense that each new available bit can be used to update the LFSR in worst case linear time. This is suitable for attackers using a known plain text attack when the number of bits available might not be known ahead of time. On a side note it is interesting that generators with low linear complexity are also undesirable for Monte Carlo and quasi-Monte Carlo based applications [66].

2.2.2 N-adic Complexity

Linear complexity measures how large an LFSR is required to generate a given sequence. While "size" is just the number of register cells in an LFSR, it also includes

[1] For a T-periodic sequence, Berlekamp-Massey algorithm takes $\mathcal{O}(T^2)$ field operations. Blackburn [5] adapted this algorithm to give an asymptotically faster algorithm with time complexity $\mathcal{O}(T(\log T)^2 \log\log T)$

the memory cells for an FCSR. In this section we present two different notions of complexity for FCSR sequences and present some known results.

The number of basic register cells in an FCSR is one less than the number of coefficients in the N-adic expansion of $q+1$ where q is the connection number. For periodic sequences we also know that the number of memory cells required is at most the $\log_N$ of the number of cells in the basic register (see [23, Theorem 7.3.2]) and thus can be ignored. But for eventually periodic sequences the number of memory cells might be more than the length of the basic register and should be counted in the size of an FCSR. We consider two related measures to estimate the minimal size of an FCSR that generates a given sequence.

Let $\mathbf{S} = (s_0, s_1, \ldots)^\infty$ be an eventually periodic N-ary sequence. Consider an FCSR with connection number $q = -1 + q_1 N + \cdots + q_m N^m$, where $q_m \neq 0$, and initial memory z that outputs this sequence.

Definition 2.4. The N-adic span $\Lambda_N(\mathbf{S})$ of an N-ary eventually periodic sequence $\mathbf{S}$ is the smallest value of

$$\Lambda = m + \max(\lfloor \log_N(\sum_{i=0}^{m} q_i) \rfloor + 1, \lfloor \log_N(|z|) \rfloor + 1) + 1$$

that occurs among all FCSRs whose output is the sequence $\mathbf{S}$.

In Definition 2.4, for a given FCSR, Λ is a bound on the number of N-ary cells needed to represent the state of the FCSR. We note that the second $+1$ in Λ is for the sign bit to accommodate negative memory values.

Definition 2.5. Let $-p/q$, where $\gcd(p, q) = 1$, be the rational number whose N-adic expansion agrees with the N-adic number associated with sequence $\mathbf{S}$. The N-adic complexity of $\mathbf{S}$ is the real number

$$\lambda_N(\mathbf{S}) = \log_N(\max(|p|, |q|)).$$

So the N-adic span of a sequence is a positive integer that counts the number of N-ary cells in the register and memory of the smallest FCSR that generates the sequence and is useful from an implementation point of view. The N-adic complexity is a real number which estimates the smallest size of the basic register of an FCSR and is more useful from a mathematical point of view.

Let $\mathbf{S} = (s_0, s_1, \ldots, s_{T-1})$ be a periodic sequence with period T and let $\mathbf{S}(N) = s_0 + s_1 N + \cdots + s_{T-1} N^{T-1}$ be the integer associated with $\mathbf{S}$. Then $-\mathbf{S}(N)/(N^T - 1)$ is a rational representation of the N-adic number corresponding to $\mathbf{S}$. The N-adic complexity of $\mathbf{S}$ is given by

$$\lambda_N(\mathbf{S}) = \log_N \left(\frac{N^T - 1}{\gcd(\mathbf{S}(N), N^T - 1)} \right).$$

For an eventually periodic sequence $\mathbf{S}$, the N-adic span and the N-adic complexity are related by

$$|(\Lambda_N(\mathbf{S}) - 2) - \lambda_N(\mathbf{S})| \leq \log_N(\lambda_N(\mathbf{S})) + 1.$$

Hence we can see that for practical purposes the N-adic complexity is a reasonable estimate for the size of an FCSR.

Let $\mathbf{S}^{rev} = (s_{T-1}, \ldots, s_0)^\infty$ be the periodic sequence formed by reversing each period of $\mathbf{S}$. Then it can be shown that the linear complexity of $\mathbf{S}$ is equal to that of $\mathbf{S}^{rev}$ using the fact that reversal commutes in polynomials (see [23, Lemma 19.2.1]). But this is not true in case of N-adic complexity. We can define the reversal of an integer with respect to its N-ary representation. Since carries in multiplications go in the opposite directions, reversal does not preserve the N-adic complexity. So given an N-ary sequence one can apply the rational approximation algorithm to find an FCSR that generates the sequence or its reversal by running the algorithm on a given segment of the sequence and its reversal. Thus a sequence $\mathbf{S}$ can be considered cryptographically strong only if $\lambda_N(\mathbf{S})$ and $\lambda_N(\mathbf{S}^{rev})$ are both high.

Definition 2.6. The symmetric N-adic complexity of a periodic N-ary sequence $\mathbf{S}$ is defined as the minimum of the N-adic complexities of $\mathbf{S}$ and $\mathbf{S}^{rev}$.

Like the Berlekamp-Massey algorithm for LFSRs, there is an adaptive rational approximation algorithm that solves the register synthesis problem for FCSR sequences. Say we have the first $t = 2\lceil \lambda_N(\mathbf{S}) \rceil + 2$ bits of a binary sequence $\mathbf{S}$. The rational approximation algorithm due to Klapper and Goresky [41] produces a pair of integers $f = (f_1, f_2)$ so that $\mathbf{S} = f_1/f_2$ and $\max(|f_1|, |f_2|)$ is minimal over all such pairs of integers in time $\mathcal{O}(t^2 \log t \log\log t)$. Once the rational representation is available, the corresponding FCSR can be constructed [23, Procedure 7.3.1]. A rational approximation algorithm for $N > 2$ was given by Xu [90] and a non-adaptive algorithm based on the Euclidean algorithm was given by Arnault et al. [4]. Because of these algorithms sequences for cryptographic purposes should not only have high linear complexity but also high N-adic complexity. In general, it is desirable to use sequences that have high complexity corresponding to different classes of generators.

2.2.3 n-th Complexity Measures

Since LFSRs and FCSRs cannot generate ultimately nonperiodic sequences it is reasonable to study the complexity of prefixes of those sequences. In practice if an attacker can recover a large prefix of the key stream the system is considered vulnerable. So every prefix of a sequence should have high complexity, since otherwise an attacker can run register synthesis algorithms on finite prefixes.

Definition 2.7 ([76])**.** Let n be a positive integer and $\mathbf{S} = (s_0, s_1, \ldots)$ be an arbitrary sequence over $\mathbb{F}_q$ of length at least n. Then the n-th linear complexity $L^n(\mathbf{S})$ is the length of the shortest LFSR whose first n terms are $s_0, s_1, ..., s_{n-1}$. The sequence $L^1(\mathbf{S}), L^2(\mathbf{S}), ...$ of integers is called the linear complexity profile of $\mathbf{S}$.

The n-th linear complexity definition can be naturally extended to n-th N-adic complexity $\lambda_N^n(\mathbf{S})$.

Rueppel [76] suggested that a cryptographically strong sequence should have high linear complexity and that the linear complexity profile should follow the line $n/2$

"closely but irregularly". For eventually nonperiodic sequences the linear complexity $L^n(\mathbf{S})$ (resp. N-adic complexity $\lambda_N^n(\mathbf{S})$) tends to infinity as n increases. So the behavior of normalized linear (resp. N-adic) complexity $L^n(\mathbf{S})/n$ (resp. $\lambda_N^n(\mathbf{S})/n$) whose values are in the range $[0, 1]$ is studied. Niederreiter [63] proved Rueppel's conjecture that for large n the normalized linear complexity is usually close to $1/2$.

Proposition 2.1. *With probability* 1 *we have*

$$\lim_{n\to\infty} \frac{L^n(\mathbf{S})}{n} = \frac{1}{2}.$$

Results on the sets of accumulation points of normalized linear and N-adic complexities were obtained respectively by Dai et al. [11] and Klapper [39].

2.2.4 Error Complexity Measures

A stream cipher is insecure if all but a few symbols of the key stream can be extracted. Hence for a cryptographically strong sequence, the complexity[2] should not decrease drastically if a few symbols are changed. If it did, an attacker could modify the known prefix of the key stream and try to decrypt the result using the shift register synthesis algorithms. If the resulting sequence differed from the actual key stream by only a few symbols, the attacker could extract most of the message. This observation gives rise to k-error linear complexity of sequences introduced by Martin and Stamp [80] based on the earlier concepts of sphere complexity and weight complexity; see [12]. The notion of error complexity was naturally extended to N-adic complexity, too.

Definition 2.8. The k-error linear complexity $L_k(\mathbf{S})$ of a periodic sequence $\mathbf{S}$ is the smallest linear complexity that can be obtained by changing (substituting) k or fewer symbols of a single period and repeating the period.

We similarly define the k-error N-adic complexity $\lambda_{N,k}(\mathbf{S})$ of a periodic N-ary sequence. We generalize this to the k-operation complexity of a periodic sequence $\mathbf{S}$, which is the smallest complexity value that can be obtained by performing k or fewer operations on a single period and repeating the modified period. An operation is an insertion, a substitution, or a deletion of a symbol. Likewise we define k-delete and k-insert complexities. The error complexity measures can also be naturally extended to finite sequences.

The minimum number of modifications (substitutions, insertions, or deletions) that can be done to decrease the complexity is an important measure for the security of a sequence. It measures the level of noise that a sequence can withstand without compromising its security. For small k an attacker can potentially do an exhaustive search over all k bit modifications of the known initial segment to find generators for approximations of the given sequence.

[2]For the rest of the document when using the word "complexity" by itself we refer to either linear complexity, N-adic complexity, or a complexity measure based on any other class of generators.

Definition 2.9. Define $\texttt{minerr}(\mathbf{S})$ as the minimum number of substitutions required to modify a single period of $\mathbf{S}$ so that the linear complexity of the modified sequence is less than that of the original sequence $\mathbf{S}$.

The notions of $\texttt{mindel}(\mathbf{S})$, $\texttt{minins}(\mathbf{S})$, and $\texttt{minoper}(\mathbf{S})$ are similarly defined for deletions, insertions, and combinations of the three operations, respectively. The corresponding N-adic analogs of these measures are defined along the same lines.

2.2.5 Joint Complexity Measures

Due to implementation ease in hardware, word based stream ciphers (see proposals DRAGON, NLS, and SSS of the ECRYPT stream cipher project [14]) are gaining prominence. The stability theory of word based stream ciphers requires the study of multisequences and the associated joint linear complexity. An m-fold multisequence $\mathbb{S} = (\mathbf{S}^0, \mathbf{S}^1, \ldots, \mathbf{S}^m)$, $m \geq 1$, over $\mathbb{F}_q$ is an m-tuple of sequences $\mathbf{S}^i$, $i = 1, \ldots, m$, over $\mathbb{F}_q$. For simplicity we assume that all the m streams have same length (finite or infinite). For notational convenience we omit the dimension m from all the notation for multisequences as it will be clear from the context.

Definition 2.10. The joint linear complexity $L(\mathbb{S})$ of an m-fold multisequence $\mathbb{S}$ is the length of the shortest LFSR that simultaneously generates its component sequences $\mathbf{S}^i$, $1 \leq i \leq m$.

In Definition 2.10 simultaneous generation means that the same LFSR, that is, having the same connection polynomial, generates all the component sequences with possibly different initial loadings. For a finite length m-fold multisequence $\mathbb{S}$, the n-th joint linear complexity $L^n(\mathbb{S})$ is the length of the shortest LFSR that simultaneously generates at least the first n terms of $\mathbf{S}^i$, $1 \leq i \leq m$. Niederreiter and Wang [69] proved the following result on the asymptotic behavior of joint linear complexity.

Proposition 2.2. *Let $\mathbb{S}$ be an m-fold multisequence. With probability 1 we have*

$$\lim_{n\to\infty} \frac{L^n(\mathbb{S})}{n} = \frac{m}{m+1}.$$

Let T be a common period of the component sequences of a periodic m-fold multisequence $\mathbb{S} = (\mathbf{S}^0, \mathbf{S}^1, \ldots, \mathbf{S}^m)$. If $f_i(x)$ is the minimal connection polynomial of $\mathbf{S}^i$ we have $L(\mathbb{S}) = \text{lcm}\ (f_1(x), \ldots, f_m(x))$. If $\mathbf{S}^i(x)$ is the polynomial corresponding to a period of $\mathbf{S}^i$ we have $L(\mathbb{S}) = T - \deg(\gcd(x^T - 1, \mathbf{S}^1(x), \ldots, \mathbf{S}^m(x)))$. Next we define another complexity measure for multisequences that is useful in further analysis.

Definition 2.11 ([58])**.** The $\mathbb{F}_q$-linear complexity $L^{\mathbb{F}_q}(\mathbf{S})$ of a sequence $\mathbf{S} = (s_0, s_1, \ldots)$ over $\mathbb{F}_{q^m}$ is the length of the shortest LFSR with taps from $\mathbb{F}_q$ that generates $\mathbf{S}$.

We can see that $L^{\mathbb{F}_q}(\mathbf{S}) \geq L(\mathbf{S})$. For periodic sequences we have the following special case when $L^{\mathbb{F}_q}(\mathbf{S}) = L(\mathbf{S})$

Proposition 2.3 ([58]). *Let $T = p^v n$ where p is the characteristic of $\mathbb{F}_q$, $v \geq 0$, and $\gcd(n, p) = 1$. Let l be the multiplicative order of q in the residue class modulo n. Then the $\mathbb{F}_q$-linear complexity and the conventional linear complexity of a T-periodic sequence over $\mathbb{F}_{q^m}$ are the same if and only if $\gcd(l, m) = 1$.*

Since the m-dimensional vector space $\mathbb{F}_q^m$ is isomorphic to the extension field $\mathbb{F}_{q^m}$ as a vector space over $\mathbb{F}_q$, an m-fold multisequence can also be identified with a single sequence over $\mathbb{F}_{q^m}$. From here on, if $\mathbb{S}$ is an m-fold multisequence over $\mathbb{F}_q$, we denote by $\mathcal{S}$ its corresponding single sequence over $\mathbb{F}_{q^m}$. We can see that $L(\mathbb{S}) = L^{\mathbb{F}_q}(\mathcal{S})$.

Meidl et al. [60] defined three error joint linear complexity measures based on substitutions for both finite length and periodic multisequences. If each component sequence of $\mathbb{S}$ is arranged in a row of a matrix, each column can be identified with an element in $\mathbb{F}_{q^m}$.

Definition 2.12. The joint k-operation $\mathbb{F}_q$-linear complexity $L_k^{\mathbb{F}_q, oper}(\mathbb{S})$ of a periodic multisequence $\mathbb{S}$ is the minimum joint linear complexity obtained by performing at most k column-operations on the multisequence. A column operation is a substitution, an insertion, or a deletion of an entire column and hence can affect up to m symbols in the multisequence.

Again, the operations can be restricted to only of one type among insertions, deletions, and substitutions. The conventional k-error linear complexity is also defined.

Definition 2.13. The joint k-error linear complexity $L_k(\mathbb{S})$ of a periodic multisequence $\mathbb{S}$ is the minimum joint linear complexity obtained by substituting at most k symbols among all the mT elements in a single period of $\mathbb{S}$.

Since allowing insertions and deletions in each component sequence may result in component sequences of different periods, we restrict the operations to substitutions in Definition 2.13. We can also see that $L_k(\mathbb{S}) \geq L_k^{\mathbb{F}_q, oper}(\mathbb{S})$.

For periodic N-ary multisequences we define joint N-adic complexity similar to joint linear complexity.

Definition 2.14. Let $\mathbb{S} = (\mathbf{S}^1, \ldots, \mathbf{S}^m)$ be an m-fold N-ary periodic multisequence of period T. Then the joint N-adic complexity $\lambda_N(\mathbb{S})$ is the $\log_N$ of the smallest integer q such that there exists an FCSR with connection number q that simultaneously generates $\mathbf{S}^1, \ldots, \mathbf{S}^m$.

Let $-p_i/q_i$, $1 \leq i \leq m$, be the reduced rational representation of $\mathbf{S}^i$. Then we have $\lambda_N(\mathbb{S}) = \log_N(\text{lcm}\ (q_1, \ldots, q_m))$.

2.3 Analysis of Sequence Complexity Measures

The motivation, definitions, and significance for LFSR and FCSR based sequence complexity measures are presented in the previous section. While a few results are already given, here we present a brief survey of some classical and recent results on the subject.

2.3.1 Problems on Sequence Complexity Measures

In this section we list a few problems pertaining to the complexity measures defined in the previous section.

(i) **Counting functions and expected values:** Counting functions are expressions for the number of sequences with a given (error) complexity value. They enable us to have a better understanding of how various complexity values are distributed. They are also used to compute expected values and variance, which in turn give insights into the average level of security that can be expected when using particular sequence families. Determining the counting function, expected value, and variance of a complexity measure over sequences of a given finite length or over all periodic sequences of a given period is thus an important problem.

(ii) **Complexity bounds:** Often it is not easy to obtain exact counts. In such cases obtaining good upper and lower bounds on a complexity measure, its expected value, and its variance can also be helpful, especially if the bounds are tight.

(iii) **Asymptotic behavior:** Determining the sets of accumulation points of values of normalized complexity for ultimately nonperiodic sequences is also useful. The set of accumulation points aids in launching distinguishing attacks on stream ciphers where a cryptanalyst tries to distinguish a key stream generated by a stream cipher from a purely random sequence. So a sequence with the set of accumulation points $[0, 1]$ can be considered most random in this sense. Also, suppose we know the set of accumulation points $[B, 1 - B]$. Then if a cryptanalyst observes that an initial segment of a sequence has normalized complexity close to B, then the next symbol in the sequence is likely one that changes the complexity so that the normalized complexity increases. Likewise, if the initial segment has normalized complexity close to $1 - B$, then the next symbol in the sequence is likely one that leaves the complexity unchanged so that the normalized complexity decreases.

(iv) **Shift register synthesis:** Designing efficient algorithms to find the smallest generator that outputs a sequence given its first few elements is very important for cryptanalytic purposes. While this problem is solved for LFSRs and FCSRs, it should be considered for all generators that are currently used or proposed.

(v) **Sequence complexity computation:** A slightly different problem is to design efficient algorithms to compute a complexity measure. This problem is different from shift register synthesis in that it only computes the value of a complexity measure and it is usually assumed that all of the sequence is given as input.

(vi) **Sequence construction:** Constructing sequences with certain desired properties or classifying sequences with certain undesirable properties is an important

problem. For cryptographic purposes we would like to construct families of sequences with high complexity and error complexity values for measures associated with different classes of generators.

The problems mentioned above can also be considered for special cases when they are hard to solve in the general setting.

2.3.2 Results on Linear Complexity and N-adic Complexity

The first results on counting functions for linear complexity were obtained by Gustavson when analyzing the time complexity of Berlekamp-Massey algorithm.

Proposition 2.4 ([25]). *Let $N^n(L)$ denote the total number of sequences over $\mathbb{F}_q$ with nth linear complexity L. Then $N^n(0) = 1$ and for $1 \leq L \leq n$,*

$$N^n(L) = (q-1)q^{\min(2L-1, 2n-2L)}.$$

From Proposition 2.4 we have $N^n(n/2) = (q-1)q^{n-1}$, implying that most of the length n finite sequences have linear complexity $n/2$.

Using Gustavson's counting function, Meidl and Niederreiter [55] derived counting functions for the number of periodic and eventually periodic sequences with a given linear complexity.

Theorem 2.5. *Let $U(L)$ and $P(L)$ denote, respectively, the number of eventually periodic and purely periodic $\mathbb{F}_q$ sequences with linear complexity L. Then we have $P(0) = 1$, $U(0) = 0$, and for $L \geq 1$*

$$P(L) = \frac{q-1}{q+1}(q^{2L} - 1)$$

and

$$U(L) = \frac{q-1}{q+1}(q^{2L-1} + 1).$$

Using Theorem 2.5 Meidl and Niederreiter [55] derived counting functions for the n-th k-error linear complexity for several values of k and obtained bounds on its expected value.

Theorem 2.6. *Let $N_k^n(L)$ denote the number of $\mathbb{F}_q$ sequences of length n with k-error linear complexity L. Let $P(L)$ be as in Theorem 2.5. Then*

(i) $N_k^n(0) = \sum_{t=0}^{k} \binom{n}{t}(q-1)^t$, *for* $1 \leq k \leq n$.

(ii) $N_k^n(1) = (q-1)^2 \sum_{t=0}^{k} \binom{n}{t}(q-1)^t + \binom{n-2}{k}(q-1)^{k+1}$, *for* $1 \leq k < (n-1)/4$.

(iii) $N_k^n(n) = 0$, *for* $1 \leq k \leq n$.

(iv) If $L \geq 1, k \geq 0$, *and* $n \geq (4k+3)L$ *then*

$$N_k^n(L) = P(L) \sum_{r=0}^{k} \binom{n}{r}(q-1)^r + (q-1)^{k+1} \sum_{t=1}^{L} \binom{n-t}{k} q^{t-1} P(L-t).$$

Niederreiter [64] extended Gustavson's result and gave a partial counting function for multisequences.

Proposition 2.7. *Let $N^{n,(m)}(L)$ denote the number of length n multisequences with joint linear complexity L. Then for $m \geq 1$ and $n \geq 2$ we have*

$$N^{n,(m)}(L) = (q^m - 1)q^{(m+1)L-1}, \quad for\ 1 \leq L \leq \frac{n}{2}.$$

An expression for the counting function when $n/2 < L \leq n$ in Proposition 2.7 was given by Wang and Niederreiter [85] and at the time of this writing, closed forms are only known for $m = 2, 3$ [85, 86].

Meidl and Niederreiter [56, 58] derived counting functions and expected values for the linear and the k-error linear complexity of periodic sequences and multisequences with certain prime periods using discrete Fourier transforms (DFTs) of sequences. Here the number of periodic sequences with period T and a given linear complexity will be less than or equal to that count for finite strings of length T. This is because the restriction that T is also the period of the sequence cuts down some of the choices which would otherwise be counted. Here we only state the important relationship between DFT and linear complexity. For the actual counting functions and further results on the topic we suggest the reader to refer to papers by Meidl and Niederreiter [56, 57, 58, 60, 63].

Definition 2.15. Let T be a positive integer with $\gcd(T, q) = 1$ and let α be a primitive T-th root of unity in some finite extension of $\mathbb{F}_q$. Then the DFT of a given T-tuple $\mathbf{S}^T = [s_0, s_1, \ldots, s_{T-1}] \in \mathbb{F}_q^T$, is the tuple $A^T = [a_0, a_1, \ldots, a_{T-1}]$ where $a_j = \sum_{i=0}^{T-1} s_i \alpha^{ij}$, for $j = 0, 1, \ldots, T-1$.

We note that the DFT transformation is one-one and has an easily computed inverse image. The next theorem is a well known result stating the connection between the linear complexity of a sequence and its DFT.

Blahut's Theorem ([6, 49])**.** Let $\gcd(T, q) = 1$. Then the linear complexity of a T-periodic sequence $\mathbf{S} = (s_0, s_1, \ldots, s_{T-1})^\infty$ over $\mathbb{F}_q$ is equal to the Hamming weight of the DFT of $[s_0, s_1, \ldots, s_{T-1}]$.

Goresky et al. [24] presented an arithmetic analog of Blahut's theorem. Their result, unlike Blahut's theorem, only gives an upper bound on the 2-adic complexity $\lambda_2(\mathbf{S})$ in terms of the number of nonzero classical Fourier coefficients of $\mathbf{S}$.

Niederreiter et al. [59, 67, 65] and Hu et al. [29] have also constructed periodic sequences with large linear complexity and large k-error linear complexity. Niederreiter and Venkateswarlu presented similar results for periodic multisequences [68]. For further results on linear complexity and related measures please see the recent survey by Niederreiter [64].

Counting functions for N-adic complexity do not exist in the literature and are comparatively difficult to derive because certain nice properties of polynomials over finite fields do not apply to N-adic representations of integers. The average behavior

of N-adic complexity for periodic sequences was investigated first by Hu and Feng [30] for $N = 2$. Later Goresky and Klapper [23] generalized for any N and showed that the expected N-adic complexity of periodic sequences of period T is in $T - \mathcal{O}(\log(T))$. Hu et al. [31] also computed the expected value of joint N-adic complexity of periodic sequences for $N = 2$. The expected value of k-error N-adic complexity is not determined yet. The N-adic analogs for most of the results on k-error linear complexity are not determined. The joint N-adic complexity of multisequences also remains a fairly unexplored concept.

3 Lower Bounds on Error Complexity Measures

We defined error linear complexity and N-adic complexity measures in Section 2.2.4 and noted that they should be as large as possible for cryptographic purposes. Several results for the k-error linear complexity of sequences were discussed in Section 2.3.2. In those results we recall that the errors are only substitutions. Linear complexities of periodic sequences obtained by inserting and deleting few symbols were also studied [82, 83]. However, similar results for the N-adic complexity do not exist in the literature.

For a T-periodic sequence $\mathbf{S}$, by $\widehat{\mathbf{S}}$ denote any periodic sequence obtained by performing up to k modifications in one period of $\mathbf{S}$ and periodically repeating the modified period.

Jiang, Dai, and Imamura [32] gave a proof that the linear complexity $L(\widehat{\mathbf{S}}) \geq T/k - L(\mathbf{S})$ in each of the following three separate cases:

(i) at most k substitutions are performed;

(ii) at most k insertions are performed; or

(iii) at most k deletions are performed.

Their analysis did not allow any combination of these operations.

From Section 2.2.4 recall that the k-operation linear complexity $L_k^{oper}(\mathbf{S})$ of a periodic sequence $\mathbf{S}$ is the smallest linear complexity obtained by performing any combination of up to k substitutions, insertions, and deletions in a single period of $\mathbf{S}$ and then repeating the period. The k-operation N-adic complexity $\lambda_{N,k}^{oper}(\mathbf{S})$ is similarly defined for a N-ary sequence $\mathbf{S}$.

In this chapter

(i) We show Jiang, Dai, and Imamura's bound should be

$$L(\widehat{\mathbf{S}}) \geq \min\left(L(\mathbf{S}), \frac{T}{k} - L(\mathbf{S})\right).$$

(ii) We prove that this bound holds for *any combination* of up to k substitutions, insertions, and deletions. That is, we do not restrict all the operations to be of the same type. Thus we derive a lower bound on the k-operation linear complexity of a periodic sequence.

(iii) We derive similar bounds for the joint linear complexity of periodic multisequences.

(iv) Using a similar approach we derive a lower bound on k-operation N-adic complexity of N-ary sequences,

$$\lambda_N(\widehat{\mathbf{S}}) > \min\left(\lambda_N(\mathbf{S}), \frac{T}{k} - \lambda_N(\mathbf{S}) - 2 - \log_N\left(\frac{2}{N-1}\right)\right).$$

A portion of these results were presented at INDOCRYPT 2007 [37] and full results appear in the journal *Cryptography and Communications* [38].

3.1 Notation for k-Operation Modification

In this section we state an auxiliary result and describe the k-operation modification of a sequence. We establish the notation we use in the later sections to obtain the lower bounds.

Let $\mathbb{F}_q$ denote the finite field with q elements, where $q = p^r, r \geq 1$, and p is prime. Let $\mathbf{S} = (s_0, s_1, \ldots, s_{T-1})^\infty$ be a T-periodic sequence over $\mathbb{F}_q$ with period $(s_0, \ldots, s_{T-1})$. Let $\mathbf{S}(x) = s_0 + s_1 x + \cdots + s_{T-1}x^{T-1}$ be the polynomial corresponding to sequence $\mathbf{S}$. Recall that the sequence $\mathbf{S}$ can be represented as the power series

$$\sum_{i \geq 0} s_i x^i = \frac{\mathbf{S}(x)}{1 - x^T} = \frac{g(x)}{f(x)}, \quad \gcd(g(x), f(x)) = 1, \quad \deg(g(x)) < \deg(f(x)). \tag{3.1}$$

Then the linear complexity of $\mathbf{S}$ is

$$L(\mathbf{S}) = \deg\left(\frac{1 - x^T}{\gcd(\mathbf{S}(x), 1 - x^T)}\right) = \deg(f(x)). \tag{3.2}$$

We can see that

$$L(\mathbf{S}) \leq T.$$

In later sections we use the following lemma to derive bounds for the linear complexity after a k-operation modification of a single period. The proof is due to Jiang et al. [32].

Lemma 3.1. *Let $C(x), D(x) \in \mathbb{F}_q[x]$ with $\deg(D(x)) < \deg(C(x))$ and $C(x) \neq 0$. Define a periodic sequence $\mathbf{A} = (a_0, a_1, \ldots)$ over $\mathbb{F}_q$ by*

$$\sum_{i \geq 0} a_i x^i = \frac{D(x)}{C(x)}.$$

Define another sequence $\widetilde{\mathbf{A}} = (\widetilde{a}_0, \widetilde{a}_1, \ldots)$ by

$$\sum_{i \geq 0} \widetilde{a}_i x^i = \frac{[H(x)D(x)] \mod C(x)}{C(x)},$$

where $H(x) \in \mathbb{F}_q[x]$. Then

$$L(\widetilde{\mathbf{A}}) \leq L(\mathbf{A}). \tag{3.3}$$

If $\gcd(C(x), H(x)) = 1$, then equality holds in equation (3.3).

Let $\mathbf{S}_{-r} = (s_{T-r}, \ldots, s_{T-1}, s_0, \ldots, s_{T-r-1})^\infty$ denote the sequence obtained by shifting one period of $\mathbf{S} = (s_0, \ldots, s_{T-1})^\infty$ to the right cyclically by r symbols and repeating this modified period. It is well known that

$$L(\mathbf{S}) = L(\mathbf{S}_{-r}), \quad 1 \leq r \leq T - 1. \tag{3.4}$$

Let $\mathbf{S}$ be the original sequence of period T and $\widehat{\mathbf{S}}$ be the sequence obtained after k operations are performed on a single period of $\mathbf{S}$. Say there are k_S substitutions, k_D deletions, and k_I insertions.

We do not allow the combination when $k_D = T$, $k_I = 0$, and $k_S = 0$ as this would amount to deleting all symbols resulting in an empty sequence. Let $S, D, I \subset \{0, \ldots, T-1\}$ be sets that denote the positions of substitutions, deletions, and insertions respectively. Substitutions and deletions are performed on the elements with indices in sets S and D respectively. Insertions occur before elements with indices in the set I. More than one element can be inserted before an element with index in I and there are $|I| = k_L$ insertion positions. Thus we have $|S| = k_S, |D| = k_D$, $|I| = k_L$,

$$k = k_S + k_D + k_I, \quad \text{and} \quad k_L \leq k_I. \tag{3.5}$$

If there are a deletion and a substitution at the same place we can remove the substitution and obtain the same modified sequence. Thus we can replace our list of k modifications by a list of $l \leq k$ modifications with no deletions and substitutions at the same place. Similarly we can replace an insertion and a deletion at the same position by a substitution of the element at that position with the element to be inserted. That is, we may assume that

$$D \cap S = D \cap I = \emptyset. \tag{3.6}$$

However, an insertion and a substitution can occur at the same position. Hence if k' is the cardinality of $S \cup D \cup I$, from equations (3.5) and (3.6) we have

$$k' = |S \cup D \cup I| \leq k_S + k_D + k_L \leq k.$$

Let $t_1, \ldots, t_{k'}$ be the list of the distinct elements of $S \cup D \cup I$ so that

$$t_1 < t_2 < \cdots < t_{k'}, \quad k' = |S \cup D \cup I|.$$

From equation (3.4), by replacing $\mathbf{S}$ by a cyclic shift $\mathbf{S}_{-r}$, $0 \leq r \leq T-1$, we can make $t_1 = 0$ and

$$T - t_{k'} = \max(t_2, t_3 - t_2, \ldots, T - t_{k'}) \geq \frac{T}{k'}. \tag{3.7}$$

So from equation (3.7) we have

$$t_{k'} \leq \frac{(k'-1)T}{k'} \leq \frac{(k-1)T}{k}. \tag{3.8}$$

3.2 Error Linear Complexity Bounds

With the notation established in the previous section, we obtain a lower bound on the linear complexity of the modified sequence. We ultimately want a bound that applies when up to k modifications are made. We first prove a lower bound assuming exactly k modifications.

Theorem 3.2. *Let* $\mathbf{S}$ *be a sequence over* $\mathbb{F}_q$ *of period* T*. Let* $\widehat{\mathbf{S}}$ *be a sequence obtained after any combination of* k *substitutions, insertions, and deletions is performed on a single period of* $\mathbf{S}$ *and repeated periodically. Then*

(i) $L(\widehat{\mathbf{S}}) \geq \min(L(\mathbf{S}), T/k - L(\mathbf{S}))$ *if the number of deletions is greater than or equal to the number of insertions.*

(ii) $L(\widehat{\mathbf{S}}) \geq \min(L(\mathbf{S}), (T+1)/k - L(\mathbf{S}))$ *if the number of deletions is less than the number of insertions.*

Proof. Let $\widehat{\mathbf{S}}(x) = \widehat{s}_0 + \widehat{s}_1 x + \cdots + \widehat{s}_{T+k_I-k_D-1} x^{T+k_I-k_D-1}$ be the polynomial corresponding to the new sequence $\widehat{\mathbf{S}} = (\widehat{s}_0, \ldots, \widehat{s}_{T+k_I-k_D-1})^\infty$ as in equation (3.1). The generating function of the new sequence is

$$\sum_{i\geq 0} \widehat{s}_i x^i = \frac{\widehat{\mathbf{S}}(x)}{1 - x^{T+k_I-k_D}}. \tag{3.9}$$

We consider two cases based on whether the number of insertions is greater than the number of deletions.

Case 1: $k_I \leq k_D$
Let

$$B(x) = x^{k_D-k_I}\widehat{\mathbf{S}}(x) - \mathbf{S}(x). \tag{3.10}$$

Since $t_{k'}$ is the position where the last operation is made, the last $T-1-t_{k'}$ coefficients are the same in $x^{k_D-k_I}\widehat{\mathbf{S}}(x)$ and $\mathbf{S}(x)$. Thus

$$\deg B(x) \leq (T-1) - (T-1-t_{k'}) = t_{k'}. \tag{3.11}$$

From equations (3.1), (3.9), and (3.10) we have

$$\begin{aligned}
\sum_{i\geq 0} \widehat{s}_i x^i &= \frac{\widehat{\mathbf{S}}(x)}{1 - x^{T+k_I-k_D}} \\
&= \frac{x^{k_I-k_D}(\mathbf{S}(x) + B(x))}{1 - x^{T+k_I-k_D}} \\
&= \frac{(g(x)(1-x^T))/f(x) + B(x)}{x^{k_D-k_I} - x^T} \\
&= \frac{g(x)(1-x^T) + f(x)B(x)}{f(x)(x^{k_D-k_I} - x^T)}.
\end{aligned}$$

Next we can apply Lemma 3.1 with $\mathbf{A} = \widehat{\mathbf{S}}$ and $H(x) = f(x)$. Hence $\widetilde{\mathbf{A}}$ is the sequence represented by

$$\begin{aligned}
\sum_{i\geq 0} \widetilde{a}_i x^i &= \frac{[f(x)(g(x)(1-x^T) + f(x)B(x))] \mod (f(x)(x^{k_D-k_I} - x^T))}{f(x)(x^{k_D-k_I} - x^T)} \\
&= \frac{[g(x)(1-x^{k_D-k_I}) + f(x)B(x)] \mod (x^{k_D-k_I} - x^T)}{x^{k_D-k_I} - x^T}.
\end{aligned} \tag{3.12}$$

Then $L(\widetilde{\mathbf{A}}) \leq L(\widehat{\mathbf{S}})$.

Since $k_D \leq t_{k'} + 1$, from equations (3.1), (3.2), and (3.11), we have

$$\deg(g(x)(1 - x^{k_D - k_I}) + f(x)B(x)) \leq L(\mathbf{S}) + t_{k'}. \tag{3.13}$$

We have the following two subcases based on the numerator in equation (3.12).

Case 1a: $[g(x)(1 - x^{k_D - k_I}) + f(x)B(x)] \not\equiv 0 \mod (x^{k_D - k_I} - x^T)$

From Lemma 3.1 and equations (3.8), (3.12), and (3.13), we have

$$\begin{aligned} L(\widehat{\mathbf{S}}) &\geq L(\widetilde{\mathbf{A}}) \\ &\geq T - \deg(g(x)(1 - x^{k_D - k_I}) + f(x)B(x)) \\ &\geq T - (L(\mathbf{S}) + t_{k'}) \\ &\geq T - L(\mathbf{S}) - \frac{(k-1)T}{k}. \end{aligned}$$

Thus we have

$$L(\widehat{\mathbf{S}}) \geq \frac{T}{k} - L(\mathbf{S}). \tag{3.14}$$

Case 1b: $[g(x)(1 - x^{k_D - k_I}) + f(x)B(x)] \equiv 0 \mod (x^{k_D - k_I} - x^T)$

If $L(\mathbf{S}) \geq T/k$, then the right hand side of equation (3.14) is at most 0 and so the result is trivial. Hence we may assume that

$$L(\mathbf{S}) < T/k. \tag{3.15}$$

Let

$$g(x)(1 - x^{k_D - k_I}) + f(x)B(x) = l(x)(x^{k_D - k_I} - x^T) \tag{3.16}$$

for some $l(x) \in \mathbb{F}_q[x]$. From equations (3.13) and (3.8) we have

$$\deg(l(x)(x^{k_D - k_I} - x^T)) \leq L(\mathbf{S}) + \frac{(k-1)T}{k}.$$

So from equation (3.15) we get $\deg(l(x)) \leq L(\mathbf{S}) - T/k < 0$. From equation (3.16) this implies that $g(x)(1 - x^{k_D - k_I}) + f(x)B(x) = 0$. Hence we have

$$B(x) = \frac{g(x)(x^{k_D - k_I} - 1)}{f(x)}. \tag{3.17}$$

From equations (3.1), (3.9), (3.10), and (3.17) we have

$$\begin{aligned} \sum_{i \geq 0} \widehat{s}_i x^i &= \frac{\widehat{\mathbf{S}}(x)}{1 - x^{T + k_I - k_D}} \\ &= \frac{B(x) + \mathbf{S}(x)}{x^{k_D - k_I} - x^T} \\ &= \frac{1}{x^{k_D - k_I} - x^T} \left(\frac{g(x)(x^{k_D - k_I} - 1)}{f(x)} + \frac{g(x)(1 - x^T)}{f(x)} \right) \\ &= \frac{g(x)}{f(x)} \\ &= \sum_{i \geq 0} s_i x^i. \end{aligned} \tag{3.18}$$

From equations (3.14) and (3.18) Case 1 of the theorem is proved.

Case 2: $k_I > k_D$
We use the result of Case 1 by switching the roles of $\mathbf{S}$ and $\widehat{\mathbf{S}}$. Let the original sequence be $\mathbf{R} = \widehat{\mathbf{S}}$. Then the new sequence $\widehat{\mathbf{R}} = \mathbf{S}$ is formed by inserting $k'_I = k_D$ symbols, deleting $k'_D = k_I$, and substituting $k'_S = k_S$ symbols. So $k'_D > k'_I$. The periods of $\mathbf{R}$ and $\widehat{\mathbf{R}}$ are $T + k_I - k_D$ and T respectively. Because $\widehat{\mathbf{R}}$ is formed by modifying $\mathbf{R}$ by deleting more symbols than those inserted, from equation (3.14) we have $L(\widehat{\mathbf{R}}) \geq \min(L(\mathbf{R}), (T + k_I - k_D)/k - L(\mathbf{R}))$.

If $\min(L(\mathbf{R}), (T+k_I-k_D)/k-L(\mathbf{R})) = L(\mathbf{R})$ and $L(\mathbf{R}) \neq (T+k_I-k_D)/k-L(\mathbf{R})$ we must have been in Case 1b for $\mathbf{R}$. We also have

$$\frac{T + k_I - k_D}{k} - L(\mathbf{R}) > L(\mathbf{R}) \geq 0. \tag{3.19}$$

From equation (3.19) and the hypothesis of Case 1b we have $\mathbf{R} = \widehat{\mathbf{R}}$.

If $\min(L(\mathbf{R}), (T + k_I - k_D)/k - L(\mathbf{R})) = (T + k_I - k_D)/k - L(\mathbf{R})$ we have

$$\begin{aligned} L(\widehat{\mathbf{R}}) &\geq \frac{T + k_I - k_D}{k} - L(\mathbf{R}) \\ &\geq \frac{T+1}{k} - L(\mathbf{R}). \end{aligned}$$

This implies $L(\mathbf{R}) \geq (T + 1)/k - L(\widehat{\mathbf{R}})$. That is,

$$L(\widehat{\mathbf{S}}) \geq \frac{T+1}{k} - L(\mathbf{S}).$$

Thus Case 2 is proved. □

Example 3.1. For a simple example of Case 1b, let $T = 10$, the sequence $\mathbf{S} = (0101010101)^\infty$, and $k = 2$. Hence $T/k - L(\mathbf{S}) = 3$, which is not a lower bound for the linear complexity of the modified sequence because we can delete any two consecutive symbols to have a sequence with linear complexity 2. Similarly we can insert two symbols and use a combination of an insertion and a deletion to obtain the same linear complexity as that of the original sequence. This shows that we must include $L(\mathbf{S})$ in our lower bound. It is this term that was missing from Jiang et al.'s lower bound [32]. Their analysis does not consider the possibility of Case 1b and hence the missing term.

Remark 3.1. We note that we can shift the sequence by one position with an insertion and a deletion by deleting the last symbol and inserting it at the beginning of the period. Hence we can leave any sequence as is up to a shift using a k-operation modification, if k is even. Even when $k \geq 3$ is odd, we can shift the sequence by $(k-3)/2$ positions using $(k-3)/2$ pairs of insertion and deletion operations. For the remaining 3 operations we look for an ab in a single period where $a, b \in \mathbb{F}_q$ such that $a \neq b$. We insert an a before a, substitute the original a by b and delete the b to leave the sequence as is up to a shift. The inclusion of $L(\mathbf{S})$ in the bound in Theorem 3.2 is also needed in view of this remark.

Corollary 3.3. *Let* $\mathbf{S}$ *be a sequence over* $\mathbb{F}_q$ *of period* T*. Then the* k*-operation linear complexity satisfies*

(i) $L_k^{oper}(\mathbf{S}) \geq \min(L(\mathbf{S}), T/k - L(\mathbf{S}))$ *if the number of deletions is greater than or equal to the number of insertions.*

(ii) $L_k^{oper}(\mathbf{S}) \geq \min(L(\mathbf{S}), (T+1)/k - L(\mathbf{S}))$ *if the number of deletions is less than the number of insertions.*

Proof. We note that the lower bound established in Theorem 3.2 is monotonically nonincreasing in k. Thus if we make $l \leq k$ modifications, the bound for exactly k modifications still applies. □

Corollary 3.4. *Let* $\mathbf{S}$ *be a sequence over* $\mathbb{F}_q$ *of period* T*. Suppose there is an* $r \in \mathbb{F}_q$ *that occurs* $t > T/2$ *times in a single period of* $\mathbf{S}$*.*

(i) If $r = 0$*, then* $L(\mathbf{S}) \leq T/(2(T-t))$ *or* $L(\mathbf{S}) \geq T/(T-t)$*.*

(ii) If $r \neq 0$*, then* $L(\mathbf{S}) \leq T/(2(T-t))$ *or* $L(\mathbf{S}) \geq T/(T-t) - 1$*.*

Proof. Assume $L(\mathbf{S}) > T/(2(T-t))$. This implies that

$$L(\mathbf{S}) > \frac{T}{T-t} - L(\mathbf{S}). \tag{3.20}$$

Let $\widehat{\mathbf{S}}$ be a sequence obtained by performing $T-t$ operations on $\mathbf{S}$ and assume that the number of deletions is greater than or equal to the number of insertions. From equation (3.20) and Corollary 3.3(i) we have

$$L(\widehat{\mathbf{S}}) \geq \frac{T}{T-t} - L(\mathbf{S}). \tag{3.21}$$

If $r = 0$, by deleting or substituting a 0 for each nonzero symbol we obtain the all-0 sequence, which has linear complexity 0. So by equation (3.21) we have $L(\mathbf{S}) \geq T/(T-t)$. So we have $L(\mathbf{S}) \leq T/(2(T-t))$ or $L(\mathbf{S}) \geq T/(T-t)$. If $r \neq 0$, by deleting or substituting an r for each symbol that is not an r we obtain the all-r sequence, which has linear complexity 1. So by equation (3.21) we have $L(\mathbf{S}) \geq T/(T-t) - 1$. So we have $L(\mathbf{S}) \leq T/(2(T-t))$ or $L(\mathbf{S}) \geq T/(T-t) - 1$. □

Remark 3.2. The results of Corollary 3.4 hold for any $r \in \mathbb{F}_q$ and the corresponding t as defined in Corollary 3.4 even if $t \leq T/2$. But the results are useful only when $t > T/2$ and there can only be one element, if any, that satisfies this condition.

Here we present a result by Kurosawa et al., which we need for our next result.

Definition 3.1. For a nonnegative integer $i = \sum_{j=0}^{d-1} i_j p^j$ with $i_j \in \{0, \ldots, p-1\}$, define

$$\texttt{Prod}(i) = \prod_{j=0}^{d-1} (i_j + 1).$$

Lemma 3.5 ([43]). *Let t_0 denote the number of occurrences of 0 in a single period of a T-periodic sequence $\mathbf{S}$ over $\mathbb{F}_q$. Let $T = p^n$ for some $n \in \mathbb{Z}^+$ where p is the characteristic of $\mathbb{F}_q$. Then the minimum number of substitutions per period required to lower the linear complexity of $\mathbf{S}$, $\mathtt{minerr}(\mathbf{S})$, satisfies the following.*

(i) $\mathtt{minerr}(\mathbf{S}) = \mathtt{Prod}(T - L(\mathbf{S}))$.

(ii) $\mathtt{minerr}(\mathbf{S}) = T - t_0$ *if and only if the minimum linear complexity achievable by performing up to* $\mathtt{minerr}(\mathbf{S})$ *substitutions on* $\mathbf{S}$ *is* 0.

(iii) If $q = 2$, $\mathtt{minerr}(\mathbf{S}) < T - \mathtt{minerr}(\mathbf{S}) = T - t_0$ *if and only if the minimum linear complexity achievable by performing up to* $\mathtt{minerr}(\mathbf{S})$ *substitutions on* $\mathbf{S}$ *is* 1.

Here we obtain a lower bound on the minimum number of operations required to obtain a sequence with linear complexity less than the original sequence without any restrictions on the period. An operation is an insertion, a deletion, or a substitution.

Corollary 3.6. *Let $\mathbf{S}$ be a not-all-zero sequence over $\mathbb{F}_q$. The minimum number of operations $\mathtt{minoper}(\mathbf{S})$ per period required to lower its linear complexity of $\mathbf{S}$ satisfies the following.*

(i) $\mathtt{minoper}(\mathbf{S}) > T/(2L(\mathbf{S}))$.

(ii) If $\mathtt{minoper}(\mathbf{S}) = T - t_0$, *where t_0 is the number of occurrences of 0 in a single period of $\mathbf{S}$, then*

$$\mathtt{minoper}(\mathbf{S}) \geq \frac{T}{L(\mathbf{S})}.$$

(iii) If $T = p^n$ for some $n \in \mathbb{Z}^+$ where p is the characteristic of $\mathbb{F}_q$, then

$$\frac{T}{2L(\mathbf{S})} < \mathtt{minoper}(\mathbf{S}) \leq \mathtt{Prod}(T - L(\mathbf{S})).$$

Proof. Let $\mathbf{S}$ and $\widehat{\mathbf{S}}$ be sequences of period T as in Corollary 3.3. After performing the necessary $k = \mathtt{minoper}(\mathbf{S})$ operations we have $L(\mathbf{S}) > L(\widehat{\mathbf{S}})$. So from Corollary 3.3(i, ii) we have

$$\begin{aligned} L(\mathbf{S}) &> L(\widehat{\mathbf{S}}) \\ &\geq \min(L(\mathbf{S}), T/k - L(\mathbf{S})). \end{aligned} \tag{3.22}$$

From equation (3.22) we have $\min(L(\mathbf{S}), T/k - L(\mathbf{S})) = T/k - L(\mathbf{S})$. Hence we have $L(\mathbf{S}) > T/k - L(\mathbf{S})$. That is, $L(\mathbf{S}) > T/2k$, which implies the bound in (i). Using Remark 3.2 and Corollary 3.4(i), (ii) follows from (i). Since $\mathtt{minoper}(\mathbf{S}) \leq \mathtt{minerr}(\mathbf{S})$, (iii) follows from (i) and Lemma 3.5(i). □

3.3 Examples

In this section we discuss the tightness of the bounds established in Theorem 3.2. Because the derivation in Theorem 3.2 does not use information about the positions and relative orders of operations, it is reasonable to investigate the tightness of those bounds.

We give non-trivial examples where the lower bounds are achieved when the least period is used in calculating them. Let V_2 denote the set of odd primes v such that 2 is a primitive root modulo v^2. We need the following results due to Meidl [52] and Han et al. [26].

Lemma 3.7 ([52]). *Let $v \in V_2$ and λ be a nonnegative integer of the form*

$$\lambda = \epsilon + (v-1)\sum_{r\in R} v^r, \quad \text{where} \quad R \subseteq \{0,\ldots,n-1\} \quad \text{and} \quad \epsilon \in \{0,1\}.$$

If $\lambda \geq (v-1)v^{n-1}$, then there exists a binary sequence with least period v^n such that the linear complexity is λ and the 1-error linear complexity is $v^n - \lambda$.

Lemma 3.8 ([26]). *For any $v \in V_2$ and $0 \leq k \leq T$, the linear complexity and hence the k-error linear complexity of a v^n-periodic binary sequence belongs to*

$$\{v^n - 1, v^n\} \cup \bigcup_{r=0}^{n-1} I_r,$$

where $I_r = \{l \in \mathbb{Z} : v^n - v^{r+1} \leq l \leq v^n - (v-1)v^r\}$.

We count the number of values greater than or equal to $(v-1)v^{n-1}$ that fall in the range specified in Lemma 3.8. From Lemma 3.7 this count gives the following result on the number of values of linear complexities for which the lower bound is achieved for v^n-periodic binary sequences.

Lemma 3.9. *For any $v \in V_2$, the number of nonnegative integers λ such that there is a binary sequence with least period v^n and linear complexity λ that achieves the lower bound in Theorem 3.2 for $k = 1$ is*

$$\frac{v^{n-1}-1}{v-1} + n + 1.$$

Next we give an infinite family of binary sequences where the lower bound is met for a single deletion and a single insertion.

Example 3.2. For a prime n, consider a 2^n-periodic binary sequence $\mathbf{S}$ with linear complexity

$$c = 2 + tn, \text{ where } \frac{2^n-2}{2n} \leq t < \frac{2^n-2}{n}. \tag{3.23}$$

Also, pick $\mathbf{S}$ so that a period (or a shift of a period) corresponds to the polynomial $\mathbf{S}(x) = x(1-x)^{2^n-c}r_1(x)r_2(x)\cdots r_t(x)$, where $r_i(x)$, $1 \leq i \leq t$, are distinct irreducible

polynomials of degree n. We note that $c > 2^{n-1}$ and hence $\mathbf{S}$ has least period 2^n. Since $x^{2^n} - x$ is the product of all monic irreducible polynomials whose degrees divide n, the number of irreducible polynomials of prime degree n in $\mathbb{F}_2[x]$ is $(2^n - 2)/n$. Now deleting the 0 at the beginning of each period results in $\widehat{\mathbf{S}}$ with period $2^n - 1$ corresponding to the polynomial $\widehat{\mathbf{S}}(x) = (1-x)^{2^n-c} r_1(x) \cdots r_t(x)$. From equation (3.23) we have $c < 2^n$. Hence

$$L(\widehat{\mathbf{S}}) = \deg\left(\frac{1 - x^{2^n-1}}{\gcd(1 - x^{2^n-1}, \widehat{\mathbf{S}}(x))}\right) = 2^n - 1 - (nt+1) = 2^n - c,$$

which achieves the lower bound in Theorem 3.2. We can also find examples for one symbol insertion by choosing $c < 2^{n-1}$ in equation (3.23) and switching the roles of $\mathbf{S}$ and $\widehat{\mathbf{S}}$.

Remark 3.3. For a 2^n-periodic binary sequence $\mathbf{S}$, $\mathtt{minerr}(\mathbf{S}) = 1$ if and only if $L(\mathbf{S}) = 2^n$. From Example 3.2 we note that there exist sequences with $\mathtt{minoper}(\mathbf{S}) = 1$ even when $2^{n-1} < L(\mathbf{S}) < 2^n$. Hence these sequences serve as examples where $\mathtt{minoper}(\mathbf{S}) < \mathtt{minerr}(\mathbf{S})$ and also achieve the lower bound in Corollary 3.6(i).

Next we use a different approach to give examples where the least period of the sequence over $\mathbb{F}_q$ is used and where a set of k substitutions yields the lower bound. With this approach we can find examples where the lower bound is achieved for nonbinary periodic sequences where the period is not necessarily a power of the field characteristic. The following result is needed for the next example. If $\mathbf{S}$ and $\mathbf{R}$ are two periodic sequences with the same period, let $d_H(\mathbf{S}, \mathbf{R})$ denote the Hamming distance between a period of $\mathbf{S}$ and a period of $\mathbf{R}$.

Lemma 3.10. *Let $\mathbf{S}$ be a sequence with least period T and minimal polynomial $f(x) \in \mathbb{F}_q[x]$ of degree n. Let $\mathbf{S}(x)$ be the polynomial corresponding to a single period as in equation (3.1). Then the sequence $\mathbf{S}$ represented by $\mathbf{S}(x^l)$, $l \in \mathbb{Z}^+$, has linear complexity nl and least period Tl. Also, if $\mathbf{R}$ is a different sequence of period T and $\mathbf{R}'$ is the sequence represented by $\mathbf{R}(x^l)$, then $d_H(\mathbf{S}, \mathbf{R}) = d_H(\mathbf{S}', \mathbf{R}')$.*

Proof. Let $\gcd(1 - x^T, \mathbf{S}(x)) = m(x)$ and $\mathbf{S}(x) = g(x)m(x)$. From equation (3.2), $1 - x^T = m(x)f(x)$ and $\gcd(f(x), g(x)) = 1$. So $\gcd(f(x^l), g(x^l)) = 1$, which implies that $\gcd(1 - x^{Tl}, \mathbf{S}(x^l)) = m(x^l)$. So

$$L(\mathbf{S}') = \deg\left(\frac{1 - x^{Tl}}{\gcd(1 - x^{Tl}, \mathbf{S}(x^l))}\right) = \deg\left(\frac{f(x^l) \cdot m(x^l)}{m(x^l)}\right) = nl.$$

We note that a single period of sequence $\mathbf{S}'$ corresponding to $\mathbf{S}(x^l)$ can be obtained by placing $l-1$ zeroes after each element in one period of $\mathbf{S}$. Hence the least period of $\mathbf{S}'$ is Tl if T is the least period of $\mathbf{S}$. For the same reason, the single period Hamming distance $d_H(\mathbf{S}, \mathbf{R}) = d_H(\mathbf{S}', \mathbf{R}')$ where $\mathbf{R}$ is a sequence with period T (which may not be its least period) and $\mathbf{R}'$ is the sequence corresponding to $\mathbf{R}(x^l)$. □

Example 3.3. Let $k = 1, q = 2$, and $T = 3$. We have $1 - x^3 = (1 + x)(1 + x + x^2)$, the factorization of $1 - x^3$ into irreducible factors in $\mathbb{F}_2[x]$. Consider the sequences $\mathbf{S}$ and $\mathbf{R}$ of least periods 3 and 1 respectively, corresponding to

$$\mathbf{S}(x) = \frac{1 - x^3}{1 + x + x^2} = 1 + x,$$

$$\mathbf{R}(x) = \frac{1 - x^3}{1 + x} = 1 + x + x^2.$$

We have $\mathbf{S} = (110)^\infty$ and $\mathbf{R} = (111)^\infty$. Here the Hamming distance of one period is $d_H(\mathbf{S}, \mathbf{R}) = 1$. It is straightforward to check that $L(\mathbf{S}) = 2$ and $L(\mathbf{R}) = 1$. From Lemma 3.10, considering $\mathbf{S}'$, $\mathbf{R}'$ corresponding to $\mathbf{S}(x^l)$, $\mathbf{R}(x^l)$, $l = 1, 2, ...$, we have $L(\mathbf{S}') = 2l$ and $L(\mathbf{R}') = l$. Also, the least period of $\mathbf{S}'$ is $3l$ and the single period Hamming distance is $d_H(\mathbf{S}', \mathbf{R}') = 1$. The lower bound from Theorem 3.2 is $(3l)/1 - L(\mathbf{S}') = 3l - 2l = l$. This can be achieved by considering $\mathbf{R}'$, which can be obtained by one modification in a single period of $\mathbf{S}'$. Also, note that $\mathbf{S}'$ achieves the lower bound in Corollary 3.6(i) since $T/(2L(\mathbf{S}')) = 3l/(4l) = 3/4$ and $d_H(\mathbf{S}', \mathbf{R}') = 1$.

Example 3.4. Let $k = 2, q = 5$, and $T = 6$. We have $1 - x^6 = 4(4 + x)(1 + x)(1 + x + x^2)(1 + 4x + x^2)$, the factorization of $1 - x^6$ into irreducible factors in $\mathbb{F}_5[x]$. Consider the sequences $\mathbf{S}$ and $\mathbf{R}$ of least periods 6 and 2 respectively, corresponding to

$$\mathbf{S}(x) = \frac{(1 - x^6)(2 + x)}{1 + 4x + x^2} = 2 + 3x + x^2 + 3x^3 + 2x^4 + 4x^5,$$

$$\mathbf{R}(x) = \frac{2(1 - x^6)}{1 + x} = 2 + 3x + 2x^2 + 3x^3 + 2x^4 + 3x^5.$$

We have $\mathbf{S} = (231324)^\infty$ and $\mathbf{R} = (232323)^\infty$. Here the Hamming distance of one period is $d_H(\mathbf{S}, \mathbf{R}) = 2$. It is straightforward to check that $L(\mathbf{S}) = 2$ and $L(\mathbf{R}) = 1$. From Lemma 3.10, considering $\mathbf{S}'$, $\mathbf{R}'$ corresponding to $\mathbf{S}(x^l)$, $\mathbf{R}(x^l)$, $l = 1, 2, ...$, we have $L(\mathbf{S}') = 2l$ and $L(\mathbf{R}') = l$. Also, the least period of $\mathbf{S}'$ is $6l$ and the single period Hamming distance is $d_H(\mathbf{S}', \mathbf{R}') = 2$. The lower bound from Theorem 3.2 is $(6l)/2 - L(\mathbf{S}') = 3l - 2l = l$. This can be achieved by considering $\mathbf{R}'$, which can be obtained by two modifications in a single period of $\mathbf{S}'$. Also, note that $\mathbf{S}'$ achieves the lower bound in Corollary 3.6(i) since $T/(2L(\mathbf{S}')) = 6l/(4l) = 3/2$ and $d_H(\mathbf{S}', \mathbf{R}') = 2$.

3.4 Joint Error Linear Complexity Bounds

In this section we show that the bounds established in Corollary 3.3 also apply for periodic multisequences over $\mathbb{F}_q$.

Let $\mathbb{S} = (\mathbf{S}^1, \ldots, \mathbf{S}^m)$ denote a periodic multisequence of period T consisting of m parallel streams of sequences $\mathbf{S}^i = (s_0^i, s_1^i, \ldots)$, $1 \le j \le m$, each of period T. Recall that an m-fold multisequence $\mathbb{S}$ can be treated as a single sequence $\mathcal{S}$ over $\mathbb{F}_{q^m}$ and the $\mathbb{F}_q$-linear complexity $L^{\mathbb{F}_q}(\mathcal{S})$ of $\mathcal{S}$ is equal to the joint linear complexity $L(\mathbb{S})$ of

$\mathbb{S}$. We also recall that $L^{\mathbb{F}_q}(\mathcal{S})$ is greater than or equal to the conventional linear complexity $L(\mathcal{S})$. From these observations we have

$$L(\mathbb{S}) = L^{\mathbb{F}_q}(\mathcal{S}) \geq L(\mathcal{S}), \tag{3.24}$$

which leads to the following result.

Theorem 3.11. *Let $\mathbb{S}$ be an m-fold multisequence over $\mathbb{F}_q$ of period T. Then we have*

(i) $L_k(\mathbb{S}) \geq L_k^{\mathbb{F}_q,oper}(\mathbb{S}) \geq \min(L(\mathbb{S}), T/k - L(\mathbb{S}))$ *if the number of column deletions is greater than or equal to the number of column insertions.*

(ii) $L_k(\mathbb{S}) \geq L_k^{\mathbb{F}_q,oper}(\mathbb{S}) \geq \min(L(\mathbb{S}), (T+1)/k - L(\mathbb{S}))$ *if the number of column deletions is less than the number of column insertions.*

Proof. Let $\mathcal{S}$ be the sequence over $\mathbf{F}_{q^m}$ corresponding to the multisequence $\mathbb{S}$ over $\mathbb{F}_q$ and let $L_k^{oper}(\mathcal{S})$ denote the usual k-operation linear complexity for a single sequence over $\mathbb{F}_{q^m}$. From Corollary 3.3(i) we have

$$L_k^{oper}(\mathcal{S}) \geq \min\left(L(\mathcal{S}), \frac{T}{k} - L(\mathcal{S})\right),$$

if the number of deletions is greater than or equal to the number of insertions. From equation (3.24) we have $L_k^{\mathbb{F}_q,oper}(\mathbb{S}) \geq L_k^{oper}(\mathcal{S})$. Hence if $\min(L(\mathcal{S}), T/k - L(\mathcal{S})) = T/k - L(\mathcal{S})$, then

$$\begin{aligned} L_k^{\mathbb{F}_q,oper}(\mathbb{S}) &\geq L_k^{oper}(\mathcal{S}) \\ &\geq \frac{T}{k} - L(\mathcal{S}) \\ &\geq \frac{T}{k} - L(\mathbb{S}). \end{aligned}$$

From Definitions 2.13 and 2.12 we know $L_k^{oper}(\mathbb{S}) \geq L_k^{\mathbb{F}_q,oper}(\mathbb{S})$. Thus Case (i) is proved. If $\min(L(\mathcal{S}), T/k - L(\mathcal{S})) = L(\mathcal{S})$, from Case 2 of Theorem 3.2 we know that the modified sequence must be the same as the original sequence. Thus the first statement of this theorem is proved. The second statement follows using a similar argument as above. □

We have the following result when we perform exactly k substitutions.

Corollary 3.12. *Let $\mathbb{S}$ be an m-fold multisequence over $\mathbb{F}_q$ of period T and $\widehat{\mathbb{S}}$ be a sequence obtained by performing exactly k substitutions among all mT elements in a single period of $\mathbb{S}$. If l is the number of component sequences with at least one substitution, then we have*

$$L(\widehat{\mathbb{S}}) \geq \min\left(L(\mathbb{S}), \frac{T}{k-l+1} - L(\mathbb{S})\right).$$

Proof. Consider the arrangement of $\mathbb{S}$ in a matrix of order $m \times T$, where each row is a period of a component sequence and each column can be identified with an element in $\mathbb{F}_{q^m}$. We can see that the joint linear complexity will not change if each of the component sequences is cyclically shifted. Thus all of the l component sequences can be shifted so that there is at least one column with l substitutions. As a result, the single sequence $\mathcal{S}$ over $\mathbb{F}_{q^m}$ and the corresponding single sequence obtained by performing up to k substitutions in the multisequence $\mathbb{S}$ differ in at most $k - l + 1$ positions. So the result follows from Theorem 3.11. □

From equation (3.24) and Theorem 3.11, using a similar argument as in Corollary 3.4, we obtain the following result.

Corollary 3.13. *Let $\mathbb{S}$ be an m-fold multisequence over $\mathbb{F}_q$ of period T and let $\mathcal{S}$ be its corresponding single sequence over $\mathbb{F}_{q^m}$. Suppose there is an $r \in \mathbb{F}_{q^m}$ that occurs $t > T/2$ times in a single period of $\mathcal{S}$.*

(i) If $r = 0$, then $L(\mathbb{S}) \leq T/2(T - t)$ or $L(\mathbb{S}) \geq T/(T - t)$.

(ii) If $r \neq 0$, then $L(\mathbb{S}) \leq T/2(T - t)$ or $L(\mathbb{S}) \geq T/(T - t) - 1$.

Next we extend the bounds obtained for $\mathtt{minoper}(\mathbf{S})$ for single sequences to multisequences. Let $\mathtt{minoper}^{\mathbb{F}_q}(\mathbb{S})$ denote the minimum value of k so that $L_k^{\mathbb{F}_q,oper}(\mathbb{S}) < L(\mathbb{S})$ and let $\mathtt{minerr}(\mathbb{S})$ denote the minimum value of k so that $L_k(\mathbb{S}) < L(\mathbb{S})$.

Lemma 3.14. *Let $\mathbb{S}$ be an m-fold multisequence over $\mathbb{F}_q$ of period $T = p^n$. Set the integer $c = \max\{L(\mathbf{S}^j) : 1 \leq j \leq m\}$ and $l = |\{j : L(\mathbf{S}^j) = c, 1 \leq j \leq m\}|$. Then*

(i) $\mathtt{minerr}(\mathbb{S}) = l \cdot \mathtt{Prod}(T - L(\mathbb{S}))$.

(ii) $\mathtt{minoper}^{\mathbb{F}_q}(\mathbb{S}) \leq l \cdot \mathtt{Prod}(T - L(\mathbb{S}))$ with equality holding when $l = 1$.

Proof. If $T = p^n$, the minimal connection polynomial of a sequence with linear complexity λ is $(1 - x)^\lambda$. Since the joint minimal connection polynomial is the LCM of the minimal connection polynomials of all the component sequences, the joint linear complexity is c. If there are l sequences with linear complexity c, to lower the joint linear complexity, the linear complexity of all of the l sequences must be lowered due to the special form of connection polynomials. Hence (i) follows. By shifting the l component sequences, k symbol substitutions among mT elements can be affected using $k - l + 1$ column substitutions when $\mathbb{S}$ is arranged in the form of a matrix of order $m \times T$. Using this observation and the inequality $\mathtt{minoper}^{\mathbb{F}_q}(\mathbb{S}) \leq \mathtt{minerr}(\mathbb{S})$, (ii) follows from (i). □

From Theorem 3.11 and Lemma 3.14, using a similar argument as in Corollary 3.6 we have the following result.

Corollary 3.15. *Let $\mathbb{S}$ be an m-fold multisequence over $\mathbb{F}_q$ of period T. We have*

(i) $\mathtt{minerr}(\mathbb{S}) \geq \mathtt{minoper}^{\mathbb{F}_q}(\mathbb{S}) > T/(2L(\mathbb{S}))$.

(ii) If $T = p^n$ for some $n \in \mathbb{Z}^+$, then

$$\frac{T}{2L(\mathbb{S})} < \texttt{minoper}^{\mathbb{F}_q}(\mathbb{S}) \leq l \cdot \texttt{Prod}(T - L(\mathbb{S})),$$

where l is as in Lemma 3.14.

3.5 Error N-adic Complexity Lower Bounds

In this section we derive non-trivial lower bounds on N-adic complexity of periodic sequences using the same approach as for linear complexity. It is interesting to note that the bounds we derive here in the N-adic case and the bounds derived in the linear complexity case differ only by a small constant even though the derivation in the former case involves additions with carry.

Let $\mathbf{S} = (s_0, \ldots, s_{T-1})^\infty$ be a T-periodic sequence over $\{0, \ldots, N-1\}$ for any $N \geq 2$. Let $\mathbf{S}(N) = s_0 + s_1 N + \cdots + s_{T-1} N^{T-1}$ be the integer corresponding to sequence $\mathbf{S}$. Thus $\mathbf{S}(N)$ is an ordinary integer and $s_0, \ldots, s_{T-1}$ are the coefficients in its N-ary expansion. Recall that the sequence $\mathbf{S}$ can be represented as the N-adic number

$$\sum_{i \geq 0} s_i N^i = -\frac{\mathbf{S}(N)}{N^T - 1} = -\frac{p}{q}, \quad \text{where} \quad \gcd(p, q) = 1 \quad \text{and} \quad 0 \leq p \leq q. \tag{3.25}$$

The N-adic complexity of $\mathbf{S}$ is

$$\lambda_N(\mathbf{S}) = \log_N \left(\frac{N^T - 1}{\gcd(\mathbf{S}(N), N^T - 1)} \right) = \log_N(q).$$

We have

$$\lambda_N(\mathbf{S}) \leq \log_N(N^T - 1).$$

We need the following lemma to derive bounds for N-adic complexity. The proof is due to Hu and Feng [30].

Lemma 3.16. *Let u and v be integers with $0 \leq u \leq v$ and $v \neq 0$. Let h be a nonzero integer and $((uh) \mod v)/v = u'/v'$ where $(uh) \mod v$ means the reduced residue of uh modulo v, and $0 \leq u' \leq v'$, $v' \neq 0$. Then*

$$\frac{v'}{\gcd(u', v')} \leq \frac{v}{\gcd(u, v)}. \tag{3.26}$$

The equality in equation (3.26) holds if and only if

$$\gcd\left(h, v/\gcd(u, v)\right) = 1.$$

From Lemma 3.1 it is straightforward to show that

$$\lambda_N(\mathbf{S}) = \lambda_N(\mathbf{S}_{-r}), \quad 1 \leq r \leq T - 1.$$

We use the notation established for Theorem 3.2 in Section 3.1.

Theorem 3.17. *Let* $\mathbf{S}$ *be a sequence over* $\{0, \ldots, N-1\}$ *of period* T *and let* $\widehat{\mathbf{S}}$ *be a sequence obtained after any combination of* k *substitutions, insertions, and deletions is performed on a single period of* $\mathbf{S}$ *and repeated periodically. Then*

(1) $\lambda_N(\widehat{\mathbf{S}}) > \min(\lambda_N(\mathbf{S}), T/k - \lambda_N(\mathbf{S}) - 2 - \log_N(2/(N-1)))$ *if the number of deletions is greater than or equal to the number of insertions.*

(2) $\lambda_N(\widehat{\mathbf{S}}) > \min(\lambda_N(\mathbf{S}), (T+1)/k - \lambda_N(\mathbf{S}) - 2 - \log_N(2/(N-1)))$ *if the number of deletions is less than the number of insertions.*

Proof. Let

$$\widehat{\mathbf{S}}(N) = \widehat{s}_0 + \widehat{s}_1 N + \cdots + \widehat{s}_{T+k_I-k_D-1} N^{T+k_I-k_D}$$

be the integer corresponding to the new sequence as in equation (3.25). Now the modified sequence $\widehat{\mathbf{S}}$ corresponds to the N-adic number

$$\sum_{i \geq 0} \widehat{s}_i N^i = -\frac{\widehat{\mathbf{S}}(N)}{N^{T+k_I-k_D} - 1}. \tag{3.27}$$

We consider two cases based on whether the number of insertions is greater than the number of deletions.

Case 1: $k_I \leq k_D$
Let

$$B(N) = N^{k_D-k_I}\widehat{\mathbf{S}}(N) - \mathbf{S}(N), \tag{3.28}$$

and

$$A(N) = s_{t_{k'}+1} N^{t_{k'}+1} + \cdots + s_{T-1} N^{T-1}$$

be the sum of the leading $T - t_{k'} - 1$ terms in the N-adic expansion of $\mathbf{S}(N)$. Set

$$f(N) = N^{k_D-k_I}\widehat{\mathbf{S}}(N) - A(N)$$

and $e(N) = \mathbf{S}(N) - A(N)$. The T coefficients in the N-ary expansion of $N^{k_D-k_I}\widehat{\mathbf{S}}(N)$ are

$$0, \ldots, 0, \widehat{s}_0, \widehat{s}_1, \ldots, \widehat{s}_{T+k_I-k_D-1},$$

where there are $k_D - k_I$ zeroes before $\widehat{s}_0$. The last $T-1-t_{k'}$ coefficients are unchanged from $\mathbf{S}$, so equal the last $T - 1 - t_{k'}$ coefficients in the N-ary expansion of $\mathbf{S}(N)$, so also of $A(N)$. Since these are all the coefficients of $A(N)$, we have

$$0 \leq f(N) \leq N^{t_{k'}+1} - N^{k_D-k_I}.$$

Also, each nonzero coefficient in the N-ary expansion of $A(N)$ is the coefficient of the same degree term of $\mathbf{S}(N)$, so that

$$0 \leq e(N) \leq N^{t_{k'}+1} - 1.$$

Thus we have

$$|B(N)| = |f(N) - e(N)| \leq \max(f(N), e(N)) \leq N^{t_{k'}+1} - 1. \tag{3.29}$$

From equations (3.25), (3.27), and (3.28) we have

$$\begin{aligned}\sum_{i\geq 0}\widehat{s}_i N^i &= -\frac{\widehat{\mathbf{S}}(N)}{N^{T+k_I-k_D}-1}\\ &= -\frac{N^{k_I-k_D}(\mathbf{S}(N)+B(N))}{N^{T+k_I-k_D}-1}\\ &= -\frac{(p(N^T-1))/q+B(N)}{N^T-N^{k_D-k_I}}.\end{aligned}$$

Let

$$-\frac{u}{v} = -\frac{p(N^T-1)/q+B(N)}{N^T-N^{k_D-k_I}},$$

where $0\leq u\leq v$, $v\neq 0$, and $\gcd(u,v)=1$. We consider the following two cases.

Case 1a: $(p(N^{k_D-k_I}-1)+qB(N))\not\equiv 0 \mod (N^T-N^{k_D-k_I})$
By Lemma 3.16, with $h=q$ we have

$$\begin{aligned}v &\geq \frac{N^T-N^{k_D-k_I}}{\gcd(N^T-N^{k_D-k_I},|p(N^{k_D-k_I}-1)+qB(N)|)}\\ &\geq \frac{N^T-N^{k_D-k_I}}{|p(N^{k_D-k_I}-1)+qB(N)|}.\end{aligned} \tag{3.30}$$

Since $k_D\leq t_{k'}+1$, from equation (3.29) we have

$$|p(N^{k_D-k_I}-1)+qB(N)|<2qN^{t_{k'}+1}. \tag{3.31}$$

From equations (3.8), (3.30), and (3.31) we have

$$\begin{aligned}\log_N(v) &> \log_N(N^T-N^{k_D-k_I})-\log_N 2-\lambda_N(\mathbf{S})-t_{k'}-1\\ &\geq \log_N(N^T-N^{T-1})-\log_N 2-\lambda_N(\mathbf{S})-\frac{(k-1)T}{k}-1\\ &\geq \frac{T}{k}+\log_N\left(\frac{N-1}{N}\right)-\log_N 2-\lambda_N(\mathbf{S})-1.\end{aligned} \tag{3.32}$$

Since $\lambda_N(\widehat{\mathbf{S}})=\max(\log_N(|u|),\log_N(|v|))$, we have

$$\lambda_N(\widehat{\mathbf{S}})>\frac{T}{k}-\lambda_N(\mathbf{S})-2-\log_N\left(\frac{2}{N-1}\right). \tag{3.33}$$

Case 1b: $(p(N^{k_D-k_I}-1)+qB(N))\equiv 0 \mod (N^T-N^{k_D-k_I})$
If $\lambda_N(\mathbf{S})+2+\log_N(2/(N-1))\geq T/k$, then the right hand side of equation (3.33) is at most 0 and so the result follows immediately. Hence we may assume that

$$\lambda_N(\mathbf{S})+2+\frac{2}{N-1}<\frac{T}{k}. \tag{3.34}$$

We have

$$p(N^{k_D-k_I}-1)+qB(N)=l(N^T-N^{k_D-k_I}), \tag{3.35}$$

for some $l \in \mathbb{N}$. From equations (3.8), (3.31), (3.34), and (3.35) we have

$$\begin{aligned}
\log_N l &\le \log_N(2qN^{t_{k'}+1}) - \log_N(N^T - N^{k_D-k_I}) \\
&\le \lambda_N(\mathbf{S}) + \log_N(2) + \frac{(k-1)T}{k} + 1 - \log_N(N^T - N^{T-1}) \\
&= \lambda_N(\mathbf{S}) + 2 + \log_N\left(\frac{2}{N-1}\right) - \frac{T}{k} \\
&< 0.
\end{aligned}$$

Thus $l = 0$. From equation (3.35) this implies that

$$p(N^{k_D-k_I}-1)+qB(N)=0. \tag{3.36}$$

From equation (3.36) using a similar derivation as in case 1b of Theorem 3.2 we can show that $\widehat{\mathbf{S}} = \mathbf{S}$. Thus Case 1 of the theorem is proved.

Case 2: $k_I > k_D$
By switching the roles of $\widehat{\mathbf{S}}$ and $\mathbf{S}$, using a similar derivation as in Case 2 of Theorem 3.2 we have

$$\lambda_N(\widehat{\mathbf{S}}) > \frac{T+1}{k} - \lambda_N(\mathbf{S}) - 2 - \log_N\left(\frac{2}{N-1}\right).$$

□

Corollary 3.18. *Let* $\mathbf{S}$ *be a sequence over* $\{0,\dots,N-1\}$ *of period* T *and let* $\lambda^{oper}_{N,k}(\mathbf{S})$ *be the* k*-operation* N*-adic complexity of* $\mathbf{S}$*. Then*

(i) $\lambda^{oper}_{N,k}(\mathbf{S}) > \min(\lambda_N(\mathbf{S}), T/k - \lambda_N(\mathbf{S}) - 2 - \log_N(2/(N-1)))$ *if the number of deletions is greater than the number of insertions.*

(ii) $\lambda^{oper}_{N,k}(\mathbf{S}) > \min(\lambda_N(\mathbf{S}), \log_N(N^T-1) - (k-1)T/k - \lambda_N(\mathbf{S}) - 1)$ *if the number of deletions is equal to the number of insertions. (With* $N = 2$*, compare to Theorem 3 in [31])*

(iii) $\lambda^{oper}_{N,k}(\mathbf{S}) > \min(\lambda_N(\mathbf{S}), (T+1)/k - \lambda_N(\mathbf{S}) - 2 - \log_N(2/(N-1)))$ *if the number of deletions is less than the number of insertions.*

Proof. Parts (i) and (iii) of the corollary follow from the same observation as in Corollary 3.3. For part (ii), considering equation (3.31) with $k_D = k_I$, we have $|p(N^{k_D-k_I}-1)+qB(N)| \le qN^{t_{k'}+1}$. So from equation (3.32) with $k_D = k_I$ we have

$$\lambda_N(\widehat{\mathbf{S}}) > \min\left(\lambda_N(\mathbf{S}), \log_N(N^T-1) - \frac{(k-1)T}{k} - \lambda_N(\mathbf{S}) - 1\right),$$

which gives the lower bound in part (ii) from the same observation as in Corollary 3.3.
□

In Corollary 3.18 we note that for $N = 2$, the term $\log_N(2/(N-1)) = 1$ and for $N > 2$, $-1 < \log_N(2/(N-1)) \leq 0$ and can be ignored in stating the bound.

Corollary 3.19. *Let* $\mathbf{S}$ *be a sequence over* $\{0, \ldots, N-1\}$ *of period* T*. Suppose there is an* $r \in \{0, \ldots, N-1\}$ *that occurs* $t > T/2$ *times in a single period of* $\mathbf{S}$*.*

(i) If $r = 0$ *or* $r = N-1$ *then*

$$\lambda_N(\mathbf{S}) \leq \frac{1}{2} \cdot \left(\log_N(N^T - 1) - \frac{(T-t-1)T}{T-t} - 1 \right)$$

or

$$\lambda_N(\mathbf{S}) > \left(\log_N(N^T - 1) - \frac{(T-t-1)T}{T-t} - 1 \right).$$

(ii) If $r \neq 0$ *and* $r \neq N-1$ *then*

$$\lambda_N(\mathbf{S}) \leq \frac{1}{2} \cdot \left(\log_N(N^T - 1) - \frac{(T-t-1)T}{T-t} - 1 \right)$$

or

$$\lambda_N(\mathbf{S}) > \left(\log_N(N^T - 1) - \frac{(T-t-1)T}{T-t} - 1 \right) - 1.$$

Proof. From Corollary 3.18(ii) using an argument similar to the one in Corollary 3.4 we obtain the result when $r = 0$. Unlike the linear complexity case, the bound does not change when $r = N-1$ since the all-$N-1$ sequence has N-adic complexity 0. But when $r \notin \{0, N-1\}$ the maximum value for the N-adic complexity of an all-r sequence, $r \in \{1, \ldots, N-2\}$, is $\log_N(N-1) < 1$. From this the bound follows using an argument similar to the one in Corollary 3.4(ii). □

Considering $\log_N(N^T - 1) - (T-t-1)T/(T-t) \approx T/(T-t)$ we note that the bounds in Corollary 3.19 are similar to the linear complexity bounds in Corollary 3.4.

Corollary 3.20. *By* $\texttt{minoper}_N(\mathbf{S})$ *denote the minimum number of operations required to decrease the* N*-adic complexity of a periodic sequence* $\mathbf{S}$*. Then,*

(i) $\texttt{minoper}_N(\mathbf{S})$ *satisfies*

$$\texttt{minoper}_N(\mathbf{S}) > \frac{T}{2\lambda_N(\mathbf{S}) + 3}.$$

(ii) If $\texttt{minoper}_N(\mathbf{S}) = T - t_0$ *or* $\texttt{minoper}_N(\mathbf{S}) = T - t_{N-1}$ *where* t_i *is the number of occurrences of* i *in* $\mathbf{S}$*, we have*

$$\texttt{minoper}_N(\mathbf{S}) > \frac{T}{\lambda_N(\mathbf{S}) + 2}.$$

Proof. We note that for each of the three cases in Corollary 3.18 the second term in the minimum is greater than or equal to $T/k - \lambda_N(\mathbf{S}) - 3$. Using this we obtain the bound in part (i) by an argument similar to the one in Corollary 3.6. Using part (i), Corollary 3.19(i) and an argument similar to the one in Corollary 3.6, we obtain the bound in part (ii). □

4 Counting Functions for 2^n-Periodic Binary Sequences

While there has been considerable research on the linear complexity and the k-error linear complexity of $\mathbb{F}_q$-sequences, for $k > 0$, counting functions for the number of sequences with given k-error linear complexity or exact formulae for the expected k-error linear complexity of a random T-periodic sequence are not known even for small k such as $k = 1$. The only exception is when T is prime and q is a primitive root modulo T in which case the only possible values for linear complexity are 0, 1, $N-1$, and N. In this case Meidl and Niederreiter [56] obtained exact formulae for the k-error linear complexity for $k > 0$. Efficient algorithms to compute the k-error linear complexity of periodic sequences also do not exist for arbitrary periods. Recently, Alecu and Sălăgean [1, 2] proposed heuristic approaches to approximate the k-error linear complexity of a given periodic binary sequence.

Given the lack of results for sequences of arbitrary periods as noted in the previous paragraph, it is reasonable to first try to understand what happens when the period is of a specific special form. Often, this process of starting from special cases leads to results of more generality or equips us with a better understanding of the original problem at hand. Sequence complexity measures for periodic sequences when the period is a power of a prime have been explored extensively. It is interesting to note that most of the results on the counting functions for the linear complexity of prime power periodic sequences are obtained by analyzing efficient algorithms that compute the linear complexity and k-error linear complexity of such sequences. Games and Chan [19] gave an efficient algorithm for computing the linear complexity of a 2^n-periodic binary sequence, which we henceforth refer to as the Games-Chan algorithm. Stamp and Martin [80] extended this algorithm to compute the k-error linear complexity of 2^n-periodic binary sequences for a fixed k. Both these algorithms use $\mathcal{O}(2^n)$ bit operations. Lauder and Patterson [44] further generalized the Games-Chan algorithm and gave an algorithm to determine the k-error linear complexity for all k using one execution with $\mathcal{O}(n2^n)$ bit operations. Meidl [53] analyzed the Games-Chan algorithm to obtain the counting function and expected values for the 1-error linear complexity of 2^n-periodic binary sequences.

The Games-Chan algorithm was generalized by Ding et al. [12] and the Stamp-Martin algorithm was generalized by Kaida et al. [33] for p^n-periodic sequences over $\mathbb{F}_q$ with characteristic p. Meidl and Venkateswarlu [61] used these extended algorithms to obtain counting functions and the expected value for the 1-error linear complexity of p^n-periodic sequences over $\mathbb{F}_p$. Efficient algorithms and counting functions for the k-error linear complexity of prime power periodic binary sequences were also explored [26, 52, 89]. Efficient algorithms for periodic sequences with periods of other special forms have also been designed [7, 54, 88].

It should be noted here that efficient algorithms to compute the N-adic complexity of prime power periodic sequences are not known; known algorithms only give upper bounds. Meidl [51] gave an analog of the (extended) Games-Chan algorithm to compute an upper bound on the 2-adic complexity of a given prime power periodic

binary sequence and Dong et al. [13] provided an analog of the Martin-Stamp algorithm to compute an upper bound on the k-error N-adic complexity of prime power periodic N-ary sequences.

In this chapter we first present some preliminaries and auxiliary results and establish notation used in the latter sections. For the main contribution of this chapter, we derive formulae for the counting functions of the k-error linear complexity of 2^n-periodic binary sequences for a few specific values of k. The particular values of k will be discussed in the later sections.

4.1 Preliminaries, Auxiliary Results, and Notation

2^n-periodic binary sequences arise in several situations in cryptography and communications. For example the counter mode of a block cipher produces such sequences. Families of sequences like m-sequences and Sidelnikov sequences [78] have periods of the form $2^n - 1$, which result in 2^n-periodic sequences with a single insertion. The linear complexities of 2^n-periodic binary sequences are also interesting from a combinatorial perspective. Patterson [70] and Etzion [15] used the linear complexities of 2^n-periodic binary sequences to construct sequences and arrays with certain window properties for use in coding and communications. Recently, Lauder and Patterson [44] and Sălăgean [77] used the k-error linear complexities of 2^n-periodic binary sequences to design efficient algorithms for decoding certain binary repeated-root cyclic codes.

In this section we establish notation and give some basic results on the linear complexity of 2^n-periodic binary sequences. We then discuss the Games-Chan algorithm that computes the linear complexity of 2^n-periodic binary sequences and derive some useful properties of such sequences. We also derive results on the effect of making a small number of changes to 2^n-periodic binary sequences with a given linear complexity.

4.1.1 Linear Complexity of 2^n-Periodic Binary Sequences

Let $\mathbf{S} = (s_0, \ldots, s_{T-1})^\infty$ be a periodic binary sequence with period T. We associate the polynomial $\mathbf{S}(x) = s_0 + s_1 x + \cdots + s_{T-1}x^{T-1}$ and the corresponding T-tuple $\mathbf{S}^{(T)} = (s_0, \ldots, s_{T-1})$ to $\mathbf{S}$. Recall that the relationship between the linear complexity $L(\mathbf{S})$ and the associated polynomial $\mathbf{S}(x)$ is given by

$$L(\mathbf{S}) = T - \deg(\gcd(x^T - 1, \mathbf{S}(x))). \tag{4.1}$$

Let $w_H(\mathbf{S})$ denote the Hamming weight of the T-tuple $\mathbf{S}^{(T)}$. For $0 \le k \le T$, the k-error linear complexity of $\mathbf{S}$ satisfies

$$L_k(\mathbf{S}) = \min_{\mathbf{E}} L(\mathbf{S} + \mathbf{E}),$$

where the minimum is taken over all T-periodic binary sequences $\mathbf{E}$ with $w_H(\mathbf{E}) \le k$. Since we consider only 2^n-periodic sequences, we have $T = 2^n$ and use the observation that

$$x^T - 1 = x^{2^n} - 1 = (x-1)^{2^n} \tag{4.2}$$

for the rest of the chapter.

Recall that $\mathtt{minerr}(\mathbf{S})$ denotes the minimum value k such that the k-error linear complexity of a 2^n-periodic sequence $\mathbf{S}$ is strictly less than its linear complexity. That is,

$$\mathtt{minerr}(\mathbf{S}) = \min\{k : L_k(\mathbf{S}) < L(\mathbf{S})\}.$$

With $q = 2$ and $T = 2^n$ in Lemma 3.5(i) we have the following result.

Lemma 4.1. *For any nonzero 2^n-periodic sequence $\mathbf{S}$, we have*

$$\mathtt{minerr}(\mathbf{S}) = 2^{w_H(2^n - L(\mathbf{S}))},$$

where $w_H(j)$, denotes the Hamming weight of the binary representation of j.

For $0 \le k \le 2^n$, let $\mathcal{N}_k(L)$ and $\mathcal{A}_k(L)$ denote, respectively, the number of and the set of 2^n-periodic binary sequences with fixed k-error linear complexity L, $0 \le L \le 2^n$. When $k = 0$ we simply use $\mathcal{A}(L)$ for sequences with a given linear complexity L and $\mathcal{N}(L) = |\mathcal{A}(L)|$.

Rueppel [76] determined the counting function of linear complexity for 2^n-periodic binary sequences. Using equations (4.1) and (4.2) it is straightforward to characterize the 2^n-periodic binary sequences with fixed linear complexity.

Lemma 4.2 ([76])**.** *The counting function for the linear complexity of 2^n-periodic binary sequences is*

$$\mathcal{N}(0) = 1 \text{ and } \mathcal{N}(L) = 2^{L-1} \text{ for } 1 \le L \le 2^n. \tag{4.3}$$

Also, $\mathcal{A}(0) = \{(0, 0, \ldots)\}$ and $\mathcal{A}(L)$, where $1 \le L \le 2^n$, is equal to the set of 2^n-periodic binary sequences $\mathbf{S}$ with the corresponding polynomials

$$\mathbf{S}(x) = (1 - x)^{2^n - L} a(x),$$

where $a(x)$ is a binary polynomial with $\deg(a(x)) \le L - 1$ and $a(1) \ne 0$.

The next proposition due to Rueppel is one of the first results on the expected values of linear complexity.

Proposition 4.3 ([76])**.** *If $T = 2^n$, then the expected linear complexity E_0 of a T-periodic binary sequence is given by*

$$E_0 = T - 1 + 2^{-T}.$$

From Proposition 4.3 we can see that the linear complexity of a random 2^n-periodic binary sequence is almost 2^n.

```
Games_Chan (S, n)
    begin
    L := 0
    for i := 0 to n − 1 do
        if S_L = S_R then
            S := S_L
        else
            S := S_L + S_R
            L := L + 2^{n−i−1}
        fi
    od
    L := L + S_0
    return L
    end
```

Figure 4.1: The Games-Chan Algorithm

4.1.2 Games-Chan Algorithm

In this section we describe the Games-Chan algorithm and list some results using its analysis.

The Games-Chan algorithm [19] is a fast algorithm for computing the linear complexity of a 2^n-periodic binary sequence. For any $\mathbf{S} \in \mathcal{A}(L)$ with period $\mathbf{S}^{(2^n)} = (s_0, \ldots, s_{2^n-1})$, denote the left and right halves of $\mathbf{S}^{(2^n)}$ by

$$\mathbf{S}_L^{(2^{n-1})} = (s_0, \ldots, s_{2^{n-1}-1}) \quad \text{and} \quad \mathbf{S}_R^{(2^{n-1})} = (s_{2^{n-1}}, \ldots, s_{2^n-1}).$$

Let $\mathbf{S}_L$ and $\mathbf{S}_R$ denote the 2^{n-1} periodic sequences

$$\mathbf{S}_L = (s_0, \ldots, s_{2^{n-1}-1})^\infty \quad \text{and} \quad \mathbf{S}_R = (s_{2^{n-1}}, \ldots, s_{2^n-1})^\infty. \tag{4.4}$$

The Games-Chan algorithm can be described as in Figure 4.1.

Correctness of the Games-Chan algorithm follows by recursively applying the following result.

Proposition 4.4. *For any 2^n-periodic binary sequence $\mathbf{S} = (s_0, \ldots, s_{2^n-1})^\infty$*

(i) If $\mathbf{S}_L^{(2^{n-1})} = \mathbf{S}_R^{(2^{n-1})}$, then $L(\mathbf{S}) = L(\mathbf{S}_L)$.

(ii) If $\mathbf{S}_L^{(2^{n-1})} \neq \mathbf{S}_R^{(2^{n-1})}$, then $L(\mathbf{S}) = 2^{n-1} + L(\mathbf{S}_L + \mathbf{S}_R)$.

Proof. When $\mathbf{S}_L^{(2^{n-1})} = \mathbf{S}_R^{(2^{n-1})}$, the result follows as $\mathbf{S} = \mathbf{S}_L$.

If $\mathbf{S}_L^{(2^{n-1})} \neq \mathbf{S}_R^{(2^{n-1})}$, then the polynomial corresponding to $\mathbf{S}$ is

$$\begin{aligned}\mathbf{S}(x) &= \mathbf{S}_L(x) + x^{2^{n-1}}\mathbf{S}_R(x) \\ &= \mathbf{S}_L(x) + \mathbf{S}_R(x) + (x^{2^{n-1}} + 1)\mathbf{S}_R(x) \\ &= \mathbf{S}_L(x) + \mathbf{S}_R(x) + (1+x)^{2^{n-1}}\mathbf{S}_R(x).\end{aligned} \tag{4.5}$$

From equations (4.1) and (4.2) we have

$$L(\mathbf{S}) = 2^n - (\text{the largest power of } (1+x) \text{ dividing } \mathbf{S}(x)). \tag{4.6}$$

We know that $\deg(\mathbf{S}_L(x) + \mathbf{S}_R(x)) < 2^{n-1}$ and so by equation (4.5) the largest power of $(1+x)$ dividing $\mathbf{S}(x)$ equals the largest power of $(1+x)$ dividing $\mathbf{S}_L(x) + \mathbf{S}_R(x)$. Let this power be k. Using equation (4.6) this implies

$$L(\mathbf{S}) = 2^n - k = 2^{n-1} + 2^{n-1} - k = 2^{n-1} + L(\mathbf{S}_L + \mathbf{S}_R).$$

Thus part (ii) is proved. □

We make some observations and establish notation we use for the rest of the chapter. We note that the **for** loop in the Games-Chan algorithm shown in Figure 4.1 is executed a total of n times to compute the linear complexity of any $\mathbf{S} \in \mathcal{A}(L)$. In the ith step, $i = 0, \ldots, n-1$, the algorithm computes the linear complexity of a 2^{n-i}-periodic binary sequence. Let $\psi^i(\mathbf{S})$, $i = 0, \ldots, n-1$, denote the first period of the 2^{n-i}-periodic binary sequence considered in the ith step of the algorithm when run with input sequence $\mathbf{S}$. Let $\psi_L^i(\mathbf{S})$ and $\psi_R^i(\mathbf{S})$ denote, respectively, the left and right halves of $\psi^i(\mathbf{S})$. Let $m^i(\mathbf{S})$ denote the total value contributed to $L(\mathbf{S})$ in the algorithm during the execution from the initial value through the i-th step of the algorithm. For any two finite binary sequences, $\mathbf{S}$ and $\mathbf{S}'$, of same length let $d_H(\mathbf{S}, \mathbf{S}')$ denote the Hamming distance between $\mathbf{S}$ and $\mathbf{S}'$. We slightly abuse the notation because we also use $d_H(\mathbf{S}, \mathbf{S}')$ to denote the Hamming distance between the first periods of $\mathbf{S}, \mathbf{S}' \in A(L)$. It is straightforward to derive the following lemma from the Games-Chan algorithm.

Lemma 4.5. *Let $\mathbf{S}$ be a 2^n-periodic binary sequence. For any t integers $r_1, \ldots, r_t$ such that $0 < r_1 < r_2 < \cdots < r_t \leq n$, we have*

$$L(\mathbf{S}) = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}) \tag{4.7}$$

if and only if

$$\psi_L^{u-1}(\mathbf{S}) = \psi_R^{u-1}(\mathbf{S}) \quad \textit{exactly when} \quad u \in \{r_1, \ldots, r_t\}. \tag{4.8}$$

Fix $r_1, \ldots, r_t$ with $0 < r_1 < \cdots < r_t \leq n$ and let $L = 2^n - (2^{n-r_1} + \cdots + 2^{n-r_t})$. For any $\mathbf{S} \in \mathcal{A}(L)$ the following properties of vectors $\psi^l(\mathbf{S})$, $0 \leq l \leq n$, hold.

P1: If $l = r_i - 1$, for some $i \in \{1, \ldots, t\}$, then $w_H(\psi^l(\mathbf{S})) = 2 \cdot w_H(\psi^{l+1}(\mathbf{S}))$.

P2: For any $l \neq r_i - 1$, for all $i \in \{1, \ldots, t\}$, we have $w_H(\psi^l(\mathbf{S})) \geq w_H(\psi^{l+1}(\mathbf{S}))$.

By $\mathcal{P}_l$, $0 \le l \le n$, denote the number of distinct possibilities, over all sequences in $\mathcal{A}(L)$, for the 2^{n-l}-vector during the l-th step such that the 2^{n-l-1}-vector during the $(l+1)$-th step is fixed. Then

P3: If $l = r_i - 1$, for some $i \in \{1, \ldots, t\}$, then $\mathcal{P}_l = 1$.

P4: For any $l \neq r_i - 1$, for all $i \in \{1, \ldots, t\}$, we have $\mathcal{P}_l = 2^{2^{n-l-1}}$.

We also use the following result in the next section.

Lemma 4.6. *Let $\mathbf{S} \in \mathcal{A}(L)$ with $L \neq 0$ represented as*

$$L(\mathbf{S}) = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}), \tag{4.9}$$

where $0 < r_1 < r_2 < \cdots < r_t \le n$. Let $\mathbf{S}' \neq \mathbf{S}$ be any other 2^n-periodic binary sequence such that $m^{l-1}(\mathbf{S}) = m^{l-1}(\mathbf{S}')$ for some $l \in \{1, \ldots, n\}$. If $d_H(\psi^l(\mathbf{S}), \psi^l(\mathbf{S}')) \neq 0$, then

$$d_H(\mathbf{S}, \mathbf{S}') \ge 2^b \cdot d_H(\psi^l(\mathbf{S}), \psi^l(\mathbf{S}')), \tag{4.10}$$

where b, $1 \le b \le t$, is the unique integer determined by the inequality $r_b \le l < r_{b+1}$. Here we take $r_0 = 0$ and $r_{t+1} = n + 1$.

Proof. Since $m^{l-1}(\mathbf{S}) = m^{l-1}(\mathbf{S}')$ and $r_b \le l < r_{b+1}$, from Lemma 4.5 during the first $l-1$ steps of the Games-Chan algorithm

$$\begin{gathered} \psi_L^{u-1}(\mathbf{S}) = \psi_R^{u-1}(\mathbf{S}) \quad \text{and} \quad \psi_L^{u-1}(\mathbf{S}') = \psi_R^{u-1}(\mathbf{S}') \\ \text{if and only if} \quad u \in \{r_1, \ldots, r_b\}. \end{gathered} \tag{4.11}$$

The algorithm has a 2^{n-l}-periodic sequence as input in the l-th step. So, let

$$d_H(\psi^l(\mathbf{S}), \psi^l(\mathbf{S}')) = g, \quad \text{for some} \quad 1 \le g \le 2^{n-l}. \tag{4.12}$$

Let p_j, $0 \le p_j \le 2^{n-l} - 1$, $j = 1, \ldots, g$, be the positions where $\psi^l(\mathbf{S})$ and $\psi^l(\mathbf{S}')$ differ. If $l - r_b > 0$, then for each p_j, $j = 1, \ldots, g$, from equation (4.11) we have either

$$\psi^{l-1}(\mathbf{S})_{p_j} = \psi^{l-1}(\mathbf{S}')_{p_j} \quad \text{and} \quad \psi^{l-1}(\mathbf{S})_{p_j + 2^{n-l}} \neq \psi^{l-1}(\mathbf{S}')_{p_j + 2^{n-l}}, \tag{4.13}$$

or

$$\psi^{l-1}(\mathbf{S})_{p_j} \neq \psi^{l-1}(\mathbf{S}')_{p_j} \quad \text{and} \quad \psi^{l-1}(\mathbf{S})_{p_j + 2^{n-l}} = \psi^{l-1}(\mathbf{S}')_{p_j + 2^{n-l}}. \tag{4.14}$$

That is, for the g positions where $\psi^l(\mathbf{S})$ and $\psi^l(\mathbf{S}')$ differ, there are at least g positions where $\psi^{l-1}(\mathbf{S})$ and $\psi^{l-1}(\mathbf{S}')$ differ. Using a similar argument we see that

$$d_H(\psi^{l-c}(\mathbf{S}), \psi^{l-c}(\mathbf{S}')) \ge g, \quad 0 \le c \le l - r_b. \tag{4.15}$$

Equation (4.15) holds trivially if $l = r_b$. We note that for each of $\mathbf{S}$ and $\mathbf{S}'$, the $(r_b - 1)$-th step has a 2^{n-r_b+1}-periodic binary sequence whose left and right halves are equal and either half gets input to the r_b-th step. Thus, from equation (4.15), we have

$$d_H(\psi^{r_b-1}(\mathbf{S}), \psi^{r_b-1}(\mathbf{S}')) \ge 2g = 2 \cdot d_H(\psi^l(\mathbf{S}), \psi^l(\mathbf{S}')). \tag{4.16}$$

Using equation (4.16), the lemma follows by induction on b. □

4.1.3 Two Symbol Changes that Retain the Linear Complexity

For a 2^n-periodic binary sequence $\mathbf{S}$ and t integers $i_1, \ldots, i_t$ such that $0 \leq i_j \leq 2^n - 1$, $j = 1, \ldots, t$, denote by $\mathbf{S}_{i_1,\ldots,i_t}$ the 2^n-periodic binary sequence with the corresponding polynomial

$$\mathbf{S}_{i_1,\ldots,i_t}(x) = \mathbf{S}(x) + x^{i_1} + \cdots + x^{i_t}.$$

We say that the sequence $\mathbf{S}_{i_1,\ldots,i_t}$ is formed by a t symbol change in $\mathbf{S}$. In this section, for specific values of L, we determine the 2 symbol changes of sequences in $\mathcal{A}(L)$ that retain the linear complexity, that is, that result in sequences in $\mathcal{A}(L)$.

Lemma 4.7. *For any sequence $\mathbf{S} \in \mathcal{A}(L)$, where $L = 2^n - 2^{n-r}$ for some $1 \leq r \leq n$, and for any integer $0 \leq i \leq 2^n - 1$, the number of sequences $\mathbf{S}_{i,j} \in \mathcal{A}(L)$, where $0 \leq j \leq 2^n - 1$ and $j \neq i$, is exactly $2^{r-1} - 1$, corresponding to all $j \in \{(i + t2^{n-r+1}) \bmod 2^n : 1 \leq t \leq 2^{r-1} - 1\}$.*

Proof. First we prove the reverse direction of the lemma. Say $j = (i+t2^{n-r+1}) \bmod 2^n$ for some $1 \leq t \leq 2^{r-1} - 1$. Let the polynomial corresponding to $\mathbf{S}$ be

$$\mathbf{S}(x) = (1 + x)^{2^{n-r}} a(x), \tag{4.17}$$

for some $a(x) \in \mathbb{F}_2[x]$ such that $\deg(a(x)) \leq 2^n - 2^{n-r} - 1$ and $a(1) = 1$. Consider the polynomial

$$x^i + x^{i+t2^{n-r+1}} = x^i(1 + x^t)^{2^{n-r+1}} = x^i(1 + x)^{2^{n-r+1}}(1 + \cdots + x^{t-1})^{2^{n-r+1}}. \tag{4.18}$$

By equations (4.17), (4.18) and the definition of linear complexity we have

$$\begin{aligned}
&L(\mathbf{S}_{i,j}) \\
&= 2^n - \deg(\gcd(1 + x^{2^n}, \mathbf{S}_{i,j}(x))) \\
&= 2^n - \deg(\gcd(1 + x^{2^n}, \mathbf{S}(x) + x^i + x^{(i+t2^{n-r+1}) \bmod 2^n})) \\
&= 2^n - \deg(\gcd(1 + x^{2^n}, \mathbf{S}(x) + x^i + x^{i+t2^{n-r+1}})) \\
&= 2^n - \deg(\gcd((1 + x)^{2^n}, (1 + x)^{2^{n-r}} a(x) + x^i(1 + x)^{2^{n-r+1}}(1 + \cdots + x^{t-1})^{2^{n-r+1}})) \\
&= 2^n - 2^{n-r} = L.
\end{aligned}$$

Now we prove the forward direction. We have $\mathbf{S}_{i,j} \in \mathcal{A}(L)$. From Lemma 4.5 we have

$$\psi_L^{r-1}(\mathbf{S}) = \psi_R^{r-1}(\mathbf{S}) \quad \text{and} \quad \psi_L^{r-1}(\mathbf{S}_{i,j}) = \psi_R^{r-1}(\mathbf{S}_{i,j}). \tag{4.19}$$

Assume $j \notin \{(i + t2^{n-r+1}) \bmod 2^n : 1 \leq t \leq 2^{r-1} - 1\}$. That is i and j are not congruent modulo 2^{n-r+1}. By the procedure of the Games-Chan algorithm, since the left and right halves are not equal during the first $(r - 2)$ steps of the algorithm for both $\mathbf{S}$ and $\mathbf{S}_{i,j}$, we have

$$d_H(\psi^{r-1}(\mathbf{S}), \psi^{r-1}(\mathbf{S}_{i,j})) = 2. \tag{4.20}$$

By equations (4.19) and (4.20) we have $d_H(\psi^r(\mathbf{S}), \psi^r(\mathbf{S}_{i,j})) = 1$. This implies that $w_H(\psi^r(\mathbf{S}))$ and $w_H(\psi^r(\mathbf{S}_{i,j}))$ cannot both be odd, which contradicts the fact that $L(\mathbf{S}) = L(\mathbf{S}_{i,j}) = 2^n - 2^{n-r}$. Thus it must be the case that $j \in \{(i+t2^{n-r+1}) \bmod 2^n : 1 \leq t \leq 2^{r-1} - 1\}$. □

The corresponding result of Lemma 4.7 when $2^n - 2^{n-r} < L < 2^n - 2^{n-r-1}$, $1 \le r \le n-2$, can be derived similarly. Also, Fu et al. [18] already obtained this result and hence we omit the proof here.

Lemma 4.8. *For any sequence* $\mathbf{S} \in \mathcal{A}(L)$*, where* $2^n - 2^{n-r} < L < 2^n - 2^{n-r-1}$ *for some* $1 \le r \le n-2$*, and for any integer* $0 \le i \le 2^n - 1$*, the number of sequences* $\mathbf{S}_{i,j} \in \mathcal{A}(L)$*, where* $0 \le j \le 2^n - 1$ *and* $j \ne i$*, is exactly* $2^r - 1$*, corresponding to all* $j \in \{(i + t2^{n-r}) \bmod 2^n : 1 \le t \le 2^r - 1\}$.

4.2 Counting Function for $k_{min}(L)$-Error Linear Complexity

By equations (4.1) and (4.2) we have the following result.

Lemma 4.9 ([18])**.** *For any* 2^n*-periodic sequence* $\mathbf{S}$*,* $L(\mathbf{S}) = 2^n$ *if and only if* $w_H(\mathbf{S})$ *is odd.*

Using Lemma 4.9 and by analyzing the Games-Chan algorithm, Meidl obtained the counting function for the 1-error linear complexity of 2^n-periodic binary sequences.

Theorem 4.10 ([53])**.** *For all integers of the form*

$$L_{r,C} = 2^n - 2^{r+1} + C, \quad 1 \le r \le n-1, \quad 1 \le C \le 2^r - 1,$$

the number $\mathcal{N}_1(L_{r,C})$ *of* 2^n*-periodic binary sequences with linear complexity* 2^n *and* 1*-error linear complexity* $L_{r,C}$ *is given by*

$$\mathcal{N}_1(L_{r,C}) = 2^{2^n - 2^{r+1} + r + C}.$$

$\mathcal{N}_1(0) = 2^n$*, and if* $L \ne 0$ *is not of the form* $L_{r,C}$*, then there is no* 2^n*-periodic binary sequence with linear complexity* 2^n *and and* 1*-error linear complexity* L.

From Theorem 4.10 the number of 2^n-periodic sequences with linear complexity 2^n and 1-error linear complexity $2^n - 3$ is 2^{2^n-2}, which shows that in general the linear complexity of a 2^n-periodic sequence with linear complexity 2^n cannot be decreased drastically by changing only one bit.

Using Lemma 4.1 we also obtain this well known result.

Lemma 4.11 ([18])**.** *For any* 2^n*-periodic sequence* $\mathbf{S}$*, if* $w_H(\mathbf{S})$ *is even then* $L_1(\mathbf{S}) = L(\mathbf{S})$*. If* $w_H(\mathbf{S})$ *is odd, then* $L_2(\mathbf{S}) = L_1(\mathbf{S}) < L(\mathbf{S}) = 2^n$.

Using Lemma 4.11 and the counting function in Theorem 4.10 we have the expected value of 1-error linear complexity.

Theorem 4.12 ([53])**.** *The expected value* E_1 *of the* 1*-error linear complexity of a random* 2^n*-periodic sequence,* $n \ge 3$*, is given by*

$$E_1 = 2^n - 3 + 2^{-2^n}(2^n + 1) - \sum_{r=2}^{n-1} 2^{-2^r + r}.$$

Fengxiang and Wenfeng [17] used Meidl's [53] approach of analyzing the Games-Chan algorithm to obtain the counting function and gave the exact expression for the expected value of the 2-error linear complexity of a 2^n-periodic binary sequence with linear complexity $2^n - 1$. In this section we perform a more rigorous analysis of the Games-Chan algorithm to enumerate all the possible values of k-error linear complexity of sequences in $\mathcal{A}(L)$ for $k = 2^{w_H(2^n-L)}$, that is when k is the minimum number of changes needed to lower the linear complexity below L. For certain sets of these values, we also derive the corresponding number of sequences in $\mathcal{A}(L)$ whose k-error linear complexity equals the values in those sets. These results were presented at the 15th annual workshop on *Selected Areas in Cryptography (SAC 2008)* [36]. For the rest of this section, by $k_{min}(L)$ denote the minimum number of changes needed to lower the linear complexity of sequences in $\mathcal{A}(L)$, that is $k_{min}(L) = 2^{w_H(2^n-L)}$.

Before we proceed to the main results, we present a useful preliminary result.

Lemma 4.13. *Consider a sequence* $\mathbf{S} \in \mathcal{A}(L)$*, where*

$$L = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}),$$

with $r_0 = 0 < r_1 < r_2 < \cdots < r_t < n+1 = r_{t+1}$ *and* $1 \le t \le n-1$*. Let* l*,* $0 \le l \le n$*, be a positive integer such that* $d_H(\psi_L^l(\mathbf{S}), \psi_R^l(\mathbf{S})) \neq 0$ *and* b *be the unique integer determined by* $r_b \le l < r_{b+1}$*. Then there exists a* 2^n*-periodic binary sequence* $\mathbf{S}'$ *such that*

$$L(\mathbf{S}') \le 2^n - (2^{n-r_1} + \cdots + 2^{n-r_b} + 2^{n-l-1}) < L(\mathbf{S}) \tag{4.21}$$

and

$$d_H(\mathbf{S}, \mathbf{S}') = 2^b d_H(\psi_L^l(\mathbf{S}), \psi_R^l(\mathbf{S})). \tag{4.22}$$

Proof. To obtain $\mathbf{S}'$ we first construct a sequence of vectors $\mathbf{S}'_{l-i}$, $i = 0, \ldots, l$, in that order, by modifying the corresponding vectors $\psi^{l-i}(\mathbf{S})$ obtained during the execution of the Games-Chan algorithm on $\mathbf{S}$. Next we give the construction.

(i) Construction of $\mathbf{S}'_l$: For each position where $\psi_L^l(\mathbf{S})$ and $\psi_R^l(\mathbf{S})$ differ, we flip a bit at that position in either $\psi_L^l(\mathbf{S})$ or $\psi_R^l(\mathbf{S})$ to make these two halves equal, which gives a 2^{n-l}-vector $\mathbf{S}'_l$ whose left and right halves are equal.

(ii) To construct $\mathbf{S}'_{l-j}$, $j = 1, \ldots, l$, we follow this recursive procedure:

  a) If $l - j \notin \{r_1 - 1, \ldots, r_t - 1\}$, then for each index u, $0 \le u < 2^{n-(l-j+1)}$, where $\psi^{l-j+1}(\mathbf{S})$ differs from $\mathbf{S}'_{l-j+1}$, flip the u-th (or the $u + 2^{n-l+j-1}$-th) bit in $\psi^{l-j}(\mathbf{S})$ to form $\mathbf{S}'_{l-j}$. (So the number of flips made in $\psi^{l-j}(\mathbf{S})$ to get $\mathbf{S}'_{l-j}$ is equal to the number of flips made in $\psi^{l-j+1}(\mathbf{S})$ to construct $\mathbf{S}'_{l-j+1}$.)

  b) If $l - j \in \{r_1 - 1, \ldots, r_t - 1\}$, then obtain $\mathbf{S}'_{l-j}$ by concatenating two copies of $\mathbf{S}'_{l-j+1}$. (So the number of flips made in $\psi^{l-j}(\mathbf{S})$ to construct $\mathbf{S}'_{l-j}$ is double the number of flips made in $\psi^{l-j+1}(\mathbf{S})$ to construct $\mathbf{S}'_{l-j+1}$.)

(iii) Obtain $\mathbf{S}'$ by taking $\mathbf{S}'^{(2^n)} = \mathbf{S}'_0$.

To show that this construction gives an $\mathbf{S}'$ that satisfies equations (4.21) and (4.22), we proceed by using strong induction on l. When $l = 0$, from the construction we see that $\mathbf{S}'$ is formed by making $d_H(\psi_L^0(\mathbf{S}), \psi_R^0(\mathbf{S}))$ many changes in $\mathbf{S}^{(2^n)}$. Since $b = 0$ when $l = 0$, equation (4.22) follows. Also step (i) of the construction implies that the left and right halves are equal in $\mathbf{S}'$. From the Games-Chan algorithm we have $L(\mathbf{S}') \leq 2^n - 2^{n-1}$ from which equation (4.21) follows.

Let the result hold for all $l \in \{0, \ldots, k-1\}$ when $d_H(\psi_L^l(\mathbf{S}), \psi_R^l(\mathbf{S})) \neq 0$. We show that the result holds for $l = k$. Let b be the unique integer such that $r_b \leq k < r_{b+1}$. We may assume $d_H(\psi_L^k(\mathbf{S}), \psi_R^k(\mathbf{S})) \neq 0$. In the recursive construction of $\mathbf{S}'_{k-j}$, $j = 1, \ldots, k$, we consider two cases.

Case 1: Let there exist an integer g, $1 \leq g \leq k$, such that $k - g \notin \{r_1 - 1, \ldots, r_t - 1\}$ and $(\mathbf{S}'_{k-g})_L = (\mathbf{S}'_{k-g})_R$. Let a be the unique integer such that $r_a \leq g < r_{a+1}$. We note that $r_a \leq r_b$ and $k - g < k$. Thus by inductive hypothesis we have

$$\begin{aligned} L(\mathbf{S}') &\leq 2^n - (2^{n-r_1} + \cdots + 2^{n-r_a} + 2^{n-(k-g)-1}) \\ &< 2^n - (2^{n-r_1} + \cdots + 2^{n-r_b} + 2^{n-k-1}) \\ &< L(\mathbf{S}). \end{aligned}$$

From inductive hypothesis and the recursive steps (ii)(a) and (ii)(b) we have

$$d_H(\mathbf{S}, \mathbf{S}') = 2^a d_H(\psi_L^{k-g}(\mathbf{S}), \psi_R^{k-g}(\mathbf{S})) = 2^a \cdot 2^{b-a} d_H(\psi_L^k(\mathbf{S}), \psi_R^k(\mathbf{S})).$$

Thus the result holds when $l = k$.

Case 2: If the integer g in Case 1 does not exist, from the construction of $\mathbf{S}'$ we have $\mathbf{S}'_{k-j} = \psi^{(k-j)}(\mathbf{S}')$ for $j = 0, \ldots, k$. Also, the behavior of Games-Chan algorithm is the same for $\mathbf{S}$ and $\mathbf{S}'$ in the first $k-1$ steps and differs for the first time in the k-th step. Equation (4.22) follows from the recursive steps (ii)(a) and (ii)(b) of the construction. Equation (4.21) follows from the procedure of the Games-Chan algorithm. □

Remark 4.1. For the rest of this section we say that $\mathbf{S}'$ in Lemma 4.13 is formed by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ and propagating the changes made to the 0-th step of the Games-Chan algorithm.

4.2.1 Expression for $k_{min}(L)$-Error Linear Complexity

In this section we analyze the structure of the Games-Chan algorithm to derive an expression to enumerate all possible values of $k_{min}(L)$-error linear complexity of sequences in $\mathcal{A}(L)$ in terms of the coefficients in the binary expansion of $2^n - L$. That is, we compute $L_{k_{min}(L)}(\mathbf{S})$ for any $\mathbf{S} \in \mathcal{A}(L)$. We handle the case when $1 < L < 2^n$ as the results are simple when $L = 0$ or 1 and we already know the results when $L = 2^n$ [53]. The following lemma generalizes a result by Fengxiang and Wenfeng [17, Lemma 2] using a similar proof.

Lemma 4.14. *For any sequence* $\mathbf{S} = (s_0, \ldots, s_{2^n-1})^\infty \in \mathcal{A}(L)$, *we have* $L \leq 2^n - 2^{n-r}$, $r = 1, \ldots, n$, *if and only if*

$$\sum_{i=0}^{2^r-1} s_{j+i\cdot 2^{n-r}} = 0 \quad \textit{for} \quad j = 0, \ldots, 2^{n-r} - 1.$$

We prove an auxiliary result that is used in the main result of this section.

Lemma 4.15. *Let* $\mathbf{S} \in \mathcal{A}(L)$ *with* $1 < L < 2^n$. *Consider the representation of* L *as*

$$L = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}), \tag{4.23}$$

where $r_0 = 0 < r_1 < r_2 < \cdots < r_t < n+1 = r_{t+1}$ *and* $1 \leq t \leq n-1$. *Let* $\mathbf{S}'$ *be any* 2^n*-periodic binary sequence such that* $d_H(\mathbf{S}, \mathbf{S}') = k_{min}(L) = 2^t$ *and* $L(\mathbf{S}') = L_{2^t}(\mathbf{S})$. *Define the two integers*

$$l_1 = \min\{i : 0 \leq i \leq n-1 \quad \textit{and} \quad m^i(\mathbf{S}') \neq m^i(\mathbf{S})\} \tag{4.24}$$

and

$$\begin{aligned} l_2 = \min\{i : 0 \leq i \leq n-1 \quad & \textit{and} \quad d_H(\psi_L^i(\mathbf{S}), \psi_R^i(\mathbf{S})) = 2^{t-j} \\ & \textit{with} \quad r_j \leq i < r_{j+1}\}. \end{aligned} \tag{4.25}$$

Then we have $l_1 = l_2$.

Proof. First let us see that l_1 and l_2 are well defined. From Lemma 4.1 we know $k_{min}(L) = 2^t$, which implies $L(\mathbf{S}') < L(\mathbf{S})$. We note that there exists at least one integer i, $0 \leq i \leq n-1$, such that $m^i(\mathbf{S}') \neq m^i(\mathbf{S})$ since otherwise $L(\mathbf{S}) = L(\mathbf{S}')$. Hence the set on the right hand side of equation (4.24) is not empty. From the procedure of the Games-Chan algorithm, and using the fact that $L(\mathbf{S}') < L(\mathbf{S})$, equation (4.24) implies

$$\psi_L^{l_1}(\mathbf{S}) \neq \psi_R^{l_1}(\mathbf{S}) \quad \text{and} \quad \psi_L^{l_1}(\mathbf{S}') = \psi_R^{l_1}(\mathbf{S}'). \tag{4.26}$$

From equation (4.26) we get

$$d_H(\psi^{l_1}(\mathbf{S}), \psi^{l_1}(\mathbf{S}')) \geq d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})). \tag{4.27}$$

Let b be the unique integer determined by the inequality $r_b \leq l_1 < r_{b+1}$. Since $\psi_L^{r_t-1}(\mathbf{S}) = \psi_R^{r_t-1}(\mathbf{S})$ and because the vectors considered during all the steps of the Games-Chan algorithm, except the last one, have nonzero Hamming weight, we have $w_H(\psi^{r_t-1}(\mathbf{S})) \geq 2$. So using properties P1 and P2 we get $w_H(\psi^{l_1+1}(\mathbf{S})) \geq 2^{t-b}$ and thus

$$d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})) \geq 2^{t-b}. \tag{4.28}$$

Now we show that $d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})) = 2^{t-b}$. If not, we have $d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})) > 2^{t-b}$ by equation (4.28). By equation (4.27) this implies

$$d_H(\psi^{l_1}(\mathbf{S}), \psi^{l_1}(\mathbf{S}')) > 2^{t-b}. \tag{4.29}$$

But from Lemma 4.6 we know $d_H(\mathbf{S}, \mathbf{S}') \geq 2^b \cdot d_H(\psi^{l_1}(\mathbf{S}), \psi^{l_1}(\mathbf{S}'))$. Since $d_H(\mathbf{S}, \mathbf{S}') = 2^t$, this implies $d_H(\psi^{l_1}(\mathbf{S}), \psi^{l_1}(\mathbf{S}')) \leq 2^{t-b}$. This contradicts the inequality in (4.29). Thus we have

$$d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})) = 2^{t-b}. \tag{4.30}$$

From equation (4.30) we know that the set on the right hand side of equation (4.25) is not empty and $l_2 \leq l_1$. By a denote the unique integer determined by the inequality $r_a \leq l_2 < r_{a+1}$. Because there are a steps before the l_2-th step where the left and right halves are equal, it is evident from equation (4.25) that altering $\psi^{l_2}(\mathbf{S})$ so that $\psi_L^{l_2}(\mathbf{S}) = \psi_R^{l_2}(\mathbf{S})$ and propagating these changes to the 0-th step of the Games-Chan algorithm will require exactly $2^a \cdot 2^{t-a} = 2^t$ changes in $\mathbf{S}^{(2^n)}$. But if $l_2 < l_1$, then by Lemma 4.13 forcing $\psi_L^{l_2}(\mathbf{S}) = \psi_R^{l_2}(\mathbf{S})$ will result in a 2^n-periodic binary sequence $\mathbf{S}''$ such that $d_H(\mathbf{S}, \mathbf{S}'') = 2^t$ and $L(\mathbf{S}'') < L(\mathbf{S}')$. This contradicts the fact that $L(\mathbf{S}') = L_{2^t}(\mathbf{S})$. Thus we have $l_2 = l_1$. □

Theorem 4.16. *Let $\mathbf{S} \in \mathcal{A}(L)$ with $1 < L < 2^n$. Consider the representation of L as*

$$L = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}), \tag{4.31}$$

where $r_0 = 0 < r_1 < r_2 < \cdots < r_t < n+1 = r_{t+1}$ and $1 \leq t \leq n-1$. Define the integer $w = \min\{i : r_i = n + i - t, 1 \leq i \leq t+1\}$. Then $L_{k_{min}(L)}(\mathbf{S})$ is 0 or is in one of the two forms

$$\begin{aligned} L_{j,l,C} \quad &:= \quad 2^n - \sum_{i=1}^{j-1} 2^{n-r_i} - 2^{n-l} + C, \quad 1 \leq j \leq w-1, \\ &\qquad r_{j-1} \leq l \leq r_j - 2, \quad \textit{and} \quad 1 \leq C \leq 2^{n-l-1} - 1, \end{aligned} \tag{4.32}$$

or

$$\begin{aligned} L_{w,l,C} \quad &:= \quad 2^n - \sum_{i=1}^{w-1} 2^{n-r_i} - 2^{n-l} + C, \\ &\qquad r_{w-1} \leq l \leq r_w - 3 \quad \textit{and} \quad 1 \leq C \leq 2^{n-l-1} - 2^{t-w+1}. \end{aligned} \tag{4.33}$$

Proof. From Lemma 4.1 and equation (4.31) we have $\mathtt{minerr}(\mathbf{S}) = k_{min}(L) = 2^t$. The sequences in $\mathcal{A}(L)$ whose 2^t-error linear complexity is 0 are those with exactly 2^t 1s per period. For any other sequence $\mathbf{S}$ in $\mathcal{A}(L)$ we show that the 2^t-error linear complexity is in one of the forms as stated in the theorem.

Define the integer l as in equation (4.25). That is

$$\begin{aligned} l = \min\{i : 0 \leq i \leq n-1 \quad &\text{and} \quad d_H(\psi_L^i(\mathbf{S}), \psi_R^i(\mathbf{S})) = 2^{t-j} \\ &\text{with} \quad r_j \leq i < r_{j+1}\}. \end{aligned} \tag{4.34}$$

We already know that the set on the right hand side of equation (4.34) is not empty due to the intermediate findings of Lemma 4.15. By b denote the unique integer determined by the inequality $r_b \leq l < r_{b+1}$. From the proof of Lemma 4.15 we know that altering $\psi^l(\mathbf{S})$ such that $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ and propagating these changes to the 0-th step of the Games-Chan algorithm will require exactly 2^t changes in $\mathbf{S}^{(2^n)}$. We also

see that it is necessary to alter $\psi^l(\mathbf{S})$ so that $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ to achieve the smallest linear complexity that can be obtained by making exactly 2^t errors in $\mathbf{S}^{(2^n)}$ since the remaining $n-l$ steps can only add a maximum of 2^{n-l-1} to the linear complexity of the modified sequence.

Note that $l \neq r_j - 1$, $j = 1, \ldots, t$, since $\psi_L^{r_j-1}(\mathbf{S}) = \psi_R^{r_j-1}(\mathbf{S})$ for $j = 1, \ldots, t$. Next we show that

$$\begin{aligned} \forall \quad i: \quad & l+1 \leq i \leq n-1, \quad \text{we have} \\ & w_H(\psi^i(\mathbf{S})) = 2^{t-j} \quad \text{with} \quad r_j \leq i < r_{j+1}. \end{aligned} \tag{4.35}$$

If equation (4.35) does not hold, then let m be any integer such that $l+1 \leq m \leq n-1$ and $w_H(\psi^m(\mathbf{S})) \neq 2^{t-a}$ where a is uniquely determined by the inequality $r_a \leq m < r_{a+1}$. Since $\psi_L^{r_t-1}(\mathbf{S}) = \psi_R^{r_t-1}(\mathbf{S})$, we have $w_H(\psi^{r_t-1}(\mathbf{S})) \geq 2$. Hence using properties P1 and P2 we get $w_H(\psi^m(\mathbf{S})) \geq 2^{t-a}$. This implies $w_H(\psi^m(\mathbf{S})) > 2^{t-a}$ since we assumed $w_H(\psi^m(\mathbf{S})) \neq 2^{t-a}$. Again, using P1 and P2, we have $w_H(\psi^{l+1}(\mathbf{S})) > 2^{a-b} \cdot 2^{t-a} = 2^{t-b}$, which contradicts the fact $d_H(\psi_L^l(\mathbf{S}), \psi_R^l(\mathbf{S})) = 2^{t-b}$. Thus $w_H(\psi^m(\mathbf{S})) = 2^{t-a}$ and so equation (4.35) holds.

To obtain the form of $L_{2^t}(\mathbf{S})$ we consider two cases based on the value of w.

Case 1: $w \leq t$

First we show that $n - r_i = t - i$ for $i = w, \ldots, t$. From the definition of w in the theorem statement we have $n - r_w = t - w$. Let z, $w < z \leq t$, be the smallest integer such that $n - r_z \neq t - z$. We have $r_w = n + w - t$ and $r_w < r_{w+1} < \cdots < r_z < \cdots < r_t \leq n$. So (1) $r_z \leq n - (t - z) = n - t + z$, and (2) $r_z \geq r_w + z - w = n + w - t + z - w = n - t + z$. So $r_z = n - t + z$, which contradicts our assumption about z. Thus we have $n - r_i = t - i$ for $i = w, \ldots, t$, which implies

$$L = 2^n - (2^{n-r_1} + \cdots + 2^{n-r_{w-1}} + 2^{t-w} + 2^{t-w-1} + \cdots + 2^0). \tag{4.36}$$

From equations (4.31), (4.36) and Lemma 4.5 this means that the left and right halves are equal from the $(r_w - 1)$-th step to $(n-1)$-th step of the execution of the Games-Chan algorithm. Using the fact that $n - r_w = t - w$, this implies that the 2^{t-w+1}-vector considered during the $(r_w - 1)$-th step,

$$\psi^{r_w-1}(\mathbf{S}) = (\psi^{r_w-1}(\mathbf{S})_0, \ldots, \psi^{r_w-1}(\mathbf{S})_{2^{t-w+1}-1}) = (1, \ldots, 1), \tag{4.37}$$

is an all-1 vector.

From the definition of w, equation (4.37) implies that $w_H(\psi^{r_w-2}(\mathbf{S})) = 2^{t-w+1}$. That is,

$$d_H(\psi_L^{r_w-3}(\mathbf{S}), \psi_R^{r_w-3}(\mathbf{S})) = 2^{t-w+1}. \tag{4.38}$$

By equation (4.38) and using the definition of l in equation (4.34), we have $l \leq r_w - 3$. We consider two cases based on the value of l.

Case 1a: $r_{w-1} \leq l \leq r_w - 3$

We first note that this case occurs only when the binary expansion of L as in equation (4.31) satisfies $r_{w-1} \leq r_w - 3$. Throughout this case we use the fact that $n - r_w = t - w$.

From the definition of l in equation (4.34) we have $d_H(\psi_L^l(\mathbf{S}), \psi_R^l(\mathbf{S})) = 2^{t-w+1}$. We already know that making 2^{t-w+1} changes in $\psi^l(\mathbf{S})$ so that $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ is necessary to achieve the smallest linear complexity possible by making $k_{min}(L) = 2^t$ changes in $\mathbf{S}^{(2^n)}$. But we have to decide for each of the 2^{t-w+1} positions where $\psi_L^l(\mathbf{S})$ and $\psi_R^l(\mathbf{S})$ differ, whether the change should be made in $\psi_L^l(\mathbf{S})$ or at the corresponding position in $\psi_R^l(\mathbf{S})$. In this case we show that there is a unique of making these 2^{t-w+1} changes so that the linear complexity of the 2^{n-l-1}-periodic binary sequence with period equal to either of the equal halves obtained by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ is as small as possible. Next we describe the unique way of making these changes. That is, we show that if $\mathbf{S}'$ is a sequence constructed by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ using the procedure in Lemma 4.13 such that it has the least linear complexity among sequences constructed by the procedure, then $\mathbf{S}'$ is unique.

So we know $\psi^{l+1}(\mathbf{S}')$ is the 2^{n-l-1}-vector obtained in the $(l+1)$-th step when running the Games-Chan algorithm with input $\mathbf{S}'$. The left and right halves of the vectors from the r_{w-1}-th step to the $(r_w - 2)$-th step of the execution of Games-Chan algorithm on $\mathbf{S}$ are not equal. From equation (4.37) $\psi^{r_w-1}(\mathbf{S})$ is a 2^{t-w+1}-vector with all 1s. Hence for all $v = r_{w-1}, r_{w-1}+1, \ldots, r_w - 2$, due to the procedure of the Games-Chan algorithm, we have

$$\sum_{j=0}^{2^{r_w-v-1}-1} \psi^v(\mathbf{S})_{i+j2^{t-w+1}} = 1 \quad \text{for} \quad i = 0, \ldots, 2^{t-w+1} - 1. \tag{4.39}$$

Let p_i, $0 \le p_i \le 2^{n-l-1} - 1$, $i = 0, \ldots, 2^{t-w+1} - 1$, be the positions where $\psi_L^l(\mathbf{S})$ and $\psi_R^l(\mathbf{S})$ differ. This means $w_H(\psi^{l+1}(\mathbf{S})) = 2^{t-w+1}$ with 1s at positions p_i, $i = 0, \ldots, 2^{t-w+1} - 1$. As equation (4.39) is valid for $v = l+1$, this implies that the mapping from the set of p_is to $\{0, \ldots, 2^{t-w+1} - 1\}$ given by $p_i \mapsto p_i \bmod 2^{t-w+1}$ is one-one and onto since otherwise $w_H(\psi^{r_w-1}(\mathbf{S})) < 2^{t-w+1}$. Hence for each p_i, $i = 0, \ldots, 2^{t-w+1} - 1$, only one of the choices, that is, changing $\psi_L^l(\mathbf{S})_{p_i}$ or $\psi_R^l(\mathbf{S})_{p_i}$ results in the 2^{n-l-1}-vector $\psi^{l+1}(\mathbf{S}')$ that satisfies

$$\sum_{j=0}^{2^{r_w-l-2}-1} \psi^{l+1}(\mathbf{S}')_{i+j2^{t-w+1}} = 0 \quad \text{for} \quad i = 0, \ldots, 2^{t-w+1} - 1. \tag{4.40}$$

We show that $\mathbf{S}'$ must satisfy equation (4.40) if it has the least linear complexity among sequences constructed by the procedure in Lemma 4.13. The contribution to $L(\mathbf{S})$ during the first $l-1$ steps of the algorithm is

$$(2^{n-1} + 2^{n-2} + \cdots + 2^{n-l}) - \sum_{i=1}^{w-1} 2^{n-r_i} = 2^n - 2^{n-l} - \sum_{i=1}^{w-1} 2^{n-r_i}.$$

Thus the 2^t-error linear complexity of $\mathbf{S}$ is of the form

$$L_{2^t}(\mathbf{S}) = 2^n - 2^{n-l} - \sum_{i=1}^{w-1} 2^{n-r_i} + C, \tag{4.41}$$

where C is the linear complexity of the 2^{n-l-1}-periodic binary sequence with period $\psi^{l+1}(\mathbf{S}')$. By Lemma 4.14 the value C in equation (4.41) satisfies

$$C = L((\psi^{l+1}(\mathbf{S}'))^{\infty}) \leq 2^{n-l-1} - 2^{t-w+1} \tag{4.42}$$

if and only if equation (4.40) holds. Thus $\mathbf{S}'$ is unique since there is a unique way of making 2^{t-w+1} changes in $\psi^l(\mathbf{S})$ to get $\mathbf{S}'$ so that equation (4.40) holds.

Also, $\psi^{l+1}(\mathbf{S}')$ is not the all-zero vector from the definition of l in equation (4.34), which implies $C \geq 1$. Thus from equations (4.41) and (4.42), $L_{2^t}(\mathbf{S})$ is in the form $L_{w,l,C}$ given in equation (4.33).

Case 1b: $r_{j-1} \leq l \leq r_j - 2$, $1 \leq j \leq w-1$
From the definition of l in equation (4.34) we have $d_H(\psi_L^l(\mathbf{S}), \psi_R^l(\mathbf{S})) = 2^{t-j+1}$. Also, by equation (4.35) we have $w_H(\psi^{r_j-1}(\mathbf{S})) = 2^{t-j+1}$. Since $j \neq w$ we have $n - r_j > t - j$ and so $\psi^{r_j-1}(\mathbf{S})$ is not an all-1 vector. More specifically if

$$G = \{g : \psi^{r_j-1}(\mathbf{S})_g = 0, g \in \{0, \ldots, 2^{n-r_j+1} - 1\}\}$$

then

$$|G| = 2^{n-r_j+1} - 2^{t-j+1}. \tag{4.43}$$

Using an argument that is similar to the one in Case 1a, we have

$$L_{2^t}(\mathbf{S}) = 2^n - 2^{n-l} - \sum_{i=1}^{j-1} 2^{n-r_i} + C, \tag{4.44}$$

where C is the linear complexity of the 2^{n-l-1}-periodic binary sequence with period $\psi^{l+1}(\mathbf{S}')$, which is equal to either of the equal halves obtained by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ such that the lowest possible linear complexity is achieved. The left and right halves of the vectors considered from the l-th step to the $(r_j - 2)$-th step are not equal. So by equation (4.43) due to the procedure of the Games-Chan algorithm we have

$$\sum_{f=0}^{2^{r_j-l-1}-1} \psi^l(\mathbf{S})_{i+f2^{n-r_j+1}} = 0 \quad \text{for} \quad i \in G \tag{4.45}$$

and

$$\sum_{f=0}^{2^{r_j-l-1}-1} \psi^l(\mathbf{S})_{i+f2^{n-r_j+1}} = 1 \quad \text{for} \quad i \in \{0, \ldots, 2^{n-r_j+1} - 1\} - G. \tag{4.46}$$

Let p_i, $0 \leq p_i \leq 2^{n-l-1} - 1$, $i = 0, \ldots, 2^{t-j+1} - 1$, be the positions where $\psi_L^l(\mathbf{S})$ and $\psi_R^l(\mathbf{S})$ differ. This means $w_H(\psi^{l+1}(\mathbf{S})) = 2^{t-j+1}$. By equations (4.45) and (4.46), this implies that the mapping from the set of p_is to $\{0, \ldots, 2^{n-r_j+1} - 1\}$ given by $p_i \mapsto p_i \bmod 2^{n-r_j+1}$ is one-one, since otherwise $w_H(\psi^{r_j-1}(\mathbf{S})) < 2^{t-j+1}$. We can see that this mapping is not onto from equation (4.43). Also, no element of G occurs as the inverse image of any element of the set $\{p_i : i = 0, \ldots, 2^{t-j+1}\}$.

We split the summation in equation (4.45) into two separate summations involving terms exclusively from $\psi_L^l(\mathbf{S})$ or $\psi_R^l(\mathbf{S})$. For each $i \in G$ we have

$$\Sigma_L(l,i) = \sum_{f=0}^{2^{r_j-l-2}-1} \psi_L^l(\mathbf{S})_{i+f2^{n-r_j+1}}$$
$$\text{and} \tag{4.47}$$
$$\Sigma_R(l,i) = \sum_{f=0}^{2^{r_j-l-2}-1} \psi_R^l(\mathbf{S})_{i+f2^{n-r_j+1}}.$$

For each $i \in G$, from equations (4.45) and (4.47) we know that $\Sigma_L(l,i)+\Sigma_R(l,i) = 0$, which implies $\Sigma_L(l,i) = \Sigma_R(l,i) = 0$ or $\Sigma_L(l,i) = \Sigma_R(l,i) = 1$. Note that none of the terms involved in the summations of equation (4.45) can be altered when forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$. Using these remarks we can see that by making appropriate changes at one of the positions p_i or $p_i + 2^{n-l-1}$, for each $i = 0, \ldots, 2^{t-j+1}$ in $\psi^l(\mathbf{S})$, we can only guarantee that $w_H(\psi^{l+1}(\mathbf{S}'))$ is even by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$. Thus the value C in equation (4.44) satisfies $1 \leq C \leq 2^{n-l-1} - 1$. Hence $L_{2^t}(\mathbf{S})$ is in the form $L_{j,l,C}$, $1 \leq j \leq w-1$, as in equation (4.32).

Case 2: $w = t + 1$
The proof in this case is similar to that for Case 1 and both forms in equations (4.32) and (4.33) are identical.

This completes the proof of the theorem. □

Remark 4.2. In Section 4.2.2, by obtaining the counting function for the number of sequences in $\mathcal{A}(L)$ with fixed $k_{min}(L)$-error linear complexity, we implicitly show that there exist sequences with $k_{min}(L)$-error linear complexity equal to each of the values listed in Theorem 4.16.

4.2.2 Counting Function

In this section we derive expressions for the number of sequences in $\mathcal{A}(L)$ with fixed $k_{min}(L)$-error linear complexity. We need some preliminary results to obtain the counting function.

Next we give a generalization of Lemma 3 in Fengxiang and Wenfeng's paper [17] using a more straightforward approach.

Lemma 4.17. *Let* $\mathbf{S} \in \mathcal{A}(L)$ *such that* $1 \leq L \leq 2^n - 2^r$ *for any* $r \in \{1, \ldots, n-1\}$. *Let* $\mathbf{S}'$ *be a* 2^n*-periodic binary sequence corresponding to the polynomial*

$$\mathbf{S}'(x) = \mathbf{S}(x) + \sum_{t=0}^{g} x^{i_t},$$

where $0 \leq g \leq 2^r - 1$ *and* $i_t \in \{0, \ldots, 2^n - 1\}$, $t = 0, \ldots, g$. *If no two* i_t*s are congruent modulo* 2^r *then we have* $L(\mathbf{S}') > L(\mathbf{S})$.

Proof. Consider the two polynomials

$$a(x) = \sum_{t=0}^{g} x^{i_t} \quad \text{and} \quad a'(x) = \sum_{t=0}^{g} x^{i_t \bmod 2^r}, \tag{4.48}$$

where the integers i_t, $t = 0, \ldots, g$, are as given in the hypothesis. We have

$$\mathbf{S}'(x) = \mathbf{S}(x) + a(x). \tag{4.49}$$

First we show that $a(x)$ is not divisible by $(1+x)^{2^r}$. Since no two i_ts are congruent modulo 2^r, the polynomial $a'(x) = \sum_{t=0}^{g} x^{i_t \bmod 2^r}$ is not equal to 0 and has degree less than 2^r. Hence we have

$$(1+x)^{2^r} \nmid a'(x). \tag{4.50}$$

Any positive integer i can be uniquely represented as $i = q \cdot 2^r + (i \bmod 2^r)$ for some integer q. We have

$$x^i + x^{i \bmod 2^r} = x^{i \bmod 2^r}(1 + x^{q2^r}) = x^{i \bmod 2^r}(1+x^q)^{2^r}.$$

Thus $x^i + x^{i \bmod 2^r}$ is divisible by $(1+x)^{2^r}$. From equation (4.48), this implies

$$(1+x)^{2^r} \mid (a(x) + a'(x)). \tag{4.51}$$

From equations (4.50) and (4.51) we know that $a(x)$ is not divisible by $(1+x)^{2^r}$. Thus we have

$$\deg(\gcd((1+x)^{2^n}, a(x))) \leq 2^r - 1. \tag{4.52}$$

Since $\mathbf{S} \in \mathcal{A}(L)$, we have $\mathbf{S}(x) = (1+x)^{2^n - L} b(x)$ for some polynomial $b(x) \in \mathbb{F}_2[x]$ such that $b(1) = 1$. Since it is given that $L \leq 2^n - 2^r$, we have

$$\deg(\gcd((1+x)^{2^n}, \mathbf{S}(x))) = 2^n - L \geq 2^r. \tag{4.53}$$

From equations (4.49), (4.52), and (4.53) we have

$$\begin{aligned} L(\mathbf{S}') &= 2^n - \deg(\gcd(1 + x^{2^n}, \mathbf{S}'(x))) \\ &= 2^n - \deg(\gcd((1+x)^{2^n}, \mathbf{S}(x) + a(x))) \\ &\geq 2^n - 2^r + 1 \\ &> L(\mathbf{S}). \end{aligned}$$

This completes the proof of the lemma. □

Theorem 4.18. *Let $\mathcal{N}_{k_{min}}(L, \mathcal{C})$ be the number of sequences in $\mathcal{A}(L)$, $1 < L < 2^n$, with fixed $k_{min}(L)$-error linear complexity $\mathcal{C}$. Let L be represented as*

$$L = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}),$$

where $r_0 = 0 < r_1 < r_2 < \cdots < r_t < n+1 = r_{t+1}$ and $1 \leq t \leq n-1$. Let $L_{j,l,C}$ be defined as in equations (4.32) and (4.33) where l satisfies equation (4.34) and let

$w = \min\{i : r_i = n+i-t, 1 \le i \le t+1\}$. *For* $1 \le j \le w$, *if* $1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}$, *then we have*

$$\mathcal{N}_{k_{min}}(L, L_{j,l,C}) = 2^{\rho(j,l,C)},$$

where

$$\rho(j,l,C) = 2^n - 2^{n-l} - \sum_{i=1}^{j-1} 2^{n-r_i} + \sum_{i=0}^{w-j-1} (r_{w-i} - r_{w-i-1} - 1)2^{t-w+i+1} + (r_j - l - 1)2^{t-j+1} + C - 1. \tag{4.54}$$

Also, $\mathcal{N}_{k_{min}}(L, 0) = 2^{\rho(1,0,1)}$ *and* $\mathcal{N}_{k_{min}}(L, \mathcal{C}) = 0$ *for all* $\mathcal{C}$ *not in the form* $L_{j,l,C}$ *as in equations* (4.32) *and* (4.33).

Proof. From equations (4.32) and (4.33) the $k_{min}(L)$-error linear complexity of $\mathbf{S} \in \mathcal{A}(L)$ is of the form

$$L_{j,l,C} = 2^n - \sum_{i=1}^{j-1} 2^{n-r_i} - 2^{n-l} + C \quad \text{for} \quad 1 \le j \le w, \tag{4.55}$$

where $r_{j-1} \le l \le r_j - 2$. (For $l = r_w - 2$, there exist no positive values for C in equation (4.33) and hence no valid values for $L_{w,l,C}$.) We determine the counting function for the number of sequences in $\mathcal{A}(L)$ with $k_{min}(L)$-error linear complexity equal to each of the values $L_{j,l,C}$ in equation (4.55) when $1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}$. From the definition of l in equation (4.34) and by equation (4.35), for any $\mathbf{S} \in \mathcal{A}(L)$ if $r_{j-1} \le l \le r_j - 2$, we know

$$w_H(\psi^{l+1}(\mathbf{S})) = w_H(\psi^{r_j-1}(\mathbf{S})) = 2^{t-j+1}. \tag{4.56}$$

We consider two cases based on the value of w.

Case 1: $w \le t$

From equation (4.37) for any $\mathbf{S} \in \mathcal{A}(L)$ the 2^{t-w+1}-vector $\psi^{r_w-1}(\mathbf{S})$ is an all-1 vector.

Let $\mathcal{D}^1(l)$ be the number of distinct 2^{n-l-1}-vectors $\psi^{l+1}(\mathbf{S})$ over all $\mathbf{S} \in \mathcal{A}(L)$ such that the 2^{n-r_w+1}-vector $\psi^{r_w-1}(\mathbf{S})$ is an all-1 vector. Let $\mathfrak{S}(\mathbf{v})$ denote the support of the vector $\mathbf{v}$. To determine $\mathcal{D}^1(l)$ we make the following observations.

(i) By equation (4.35) it is evident that during the execution of Games-Chan algorithm from the $(l+1)$-th step to the $(n-1)$-th step the Hamming weight of the vectors considered does not change between two consecutive steps except when going from the $(r_i - 1)$-th step to the r_i-th step for $i = j, \ldots, t$.

(ii) Let a be an integer so that $l + 1 \le a < r_j$ or $r_i \le a \le r_{i+1} - 2$ for some $i \in \{j, \ldots, t\}$. Over all $\mathbf{S} \in \mathcal{A}(L)$ we determine the number of distinct vectors in the a-th step that result in a fixed vector $\mathbf{v}$ in the $(a+1)$-th step. First note that for any two same size binary vectors $\mathbf{x}$ and $\mathbf{y}$, the only way we can have $w_H(\mathbf{x} \oplus \mathbf{y}) = w_H(\mathbf{x}) + w_H(\mathbf{y})$ is if $\mathfrak{S}(\mathbf{x}) \cap \mathfrak{S}(\mathbf{y}) = \emptyset$. Using (i), the procedure of

the Games-Chan algorithm implies that the number of distinct vectors in the a-th step that result in a fixed vector $\mathbf{v}$ in the $(a+1)$-th step is equal to

$$\begin{aligned}&|\{(\mathbf{x},\mathbf{y}) : \mathbf{x}\oplus\mathbf{y}=\mathbf{v} \quad \text{and} \quad w_H(\mathbf{x}\oplus\mathbf{y}) = w_H(\mathbf{x}) + w_H(\mathbf{y})\}|\\ &= |\{U \subseteq \{1,\ldots,w_H(\mathbf{v})\}\}|\\ &= 2^{w_H(\mathbf{v})}.\end{aligned}$$

We already know that for any $\mathbf{S} \in \mathcal{A}(L)$ the 2^{t-w+1}-vector $\psi^{r_w-1}(\mathbf{S})$ is an all-1 vector. Also, by property P1 the Hamming weight of the vector in the (r_i-1)-th step is twice the weight of the vector in the r_i-th step, for $i = 1,\ldots,t$, in the Games-Chan algorithm. By equation (4.56) and recursively applying the observations (i), (ii), and the property P1 we obtain

$$\mathcal{D}^1(l) = \left(\prod_{i=0}^{w-j-1} (2^{2^{t-w+i+1}})^{(r_{w-i}-r_{w-i-1}-1)}\right)(2^{r_j-l-2})^{2^{t-j+1}}. \tag{4.57}$$

Let $\mathbf{S}'$ be the sequence obtained by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ in the execution of the Games-Chan algorithm on $\mathbf{S}$ using 2^{t-j+1} changes and propagating the changes made to the 0-th step such that the least linear complexity is achieved by making $k_{min}(L)$ errors in $\mathbf{S}^{(2^n)}$. So $\psi^{l+1}(\mathbf{S}')$ is the vector obtained in this process when forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ in the l-th step. Let $\mathcal{D}^2(C)$, $1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}$, be the number of choices for $\psi^{l+1}(\mathbf{S}')$ such that the linear complexity of the 2^{n-l-1}-periodic sequence with period $\psi^{l+1}(\mathbf{S}')$ is C. By equation (4.3), we have

$$\mathcal{D}^2(C) = 2^{C-1} \quad \text{for} \quad 1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}. \tag{4.58}$$

Over all $\mathbf{S} \in \mathcal{A}(L)$, for a fixed $\psi^{l+1}(\mathbf{S}) = \mathbf{v}$ with $w_H(\mathbf{v}) = 2^{t-j+1}$ and for a fixed choice of $\psi^{l+1}(\mathbf{S}')$ with $L((\psi^{l+1}(\mathbf{S}'))^\infty) = C$, the number of possibilities, denoted by $\mathcal{D}^3(l)$, for $\psi^l(\mathbf{S})$ such that $\psi_L^l(\mathbf{S}) + \psi_R^l(\mathbf{S}) = \mathbf{v}$ and $d_H(\psi^l(\mathbf{S}), \psi^{l+1}(\mathbf{S}') \mid \psi^{l+1}(\mathbf{S}')) = 2^{t-j+1}$ is

$$\mathcal{D}^3(l) = 2^{2^{t-j+1}}, \tag{4.59}$$

where $\psi^{l+1}(\mathbf{S}') \mid \psi^{l+1}(\mathbf{S}')$ is the 2^{n-l}-vector formed by concatenating two copies of $\psi^{l+1}(\mathbf{S}')$.

Let p_i, $0 \le p_i \le 2^{n-l-1} - 1$, $i = 0,\ldots,2^{t-j+1}-1$, be the positions where $\psi_L^l(\mathbf{S})$ and $\psi_R^l(\mathbf{S})$ differ. In Cases 1a and 1b of the proof of Theorem 4.16, the mapping from the set of p_is to $\{0,\ldots,2^{n-r_j+1}-1\}$ given by $p_i \mapsto p_i \bmod 2^{n-r_j+1}$ is one-one. Using this mapping and the condition $1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}$, by Lemma 4.17 for fixed $\psi^{l+1}(\mathbf{S})$ and $\psi^{l+1}(\mathbf{S}')$ each of the $2^{2^{t-j+1}}$ possibilities for $\psi^l(\mathbf{S})$ satisfies

$$L(\psi_L^l(\mathbf{S})) > C \quad \text{and} \quad L(\psi_R^l(\mathbf{S})) > C. \tag{4.60}$$

By equations (4.57)–(4.60), using properties P3 and P4 recursively we obtain

$$\mathcal{N}_{k_{min}}(L, L_{j,l,C}) = \mathcal{P}_0\mathcal{P}_1\cdots\mathcal{P}_{l-1}\mathcal{D}^1(l)\mathcal{D}^2(C)\mathcal{D}^3(l). \tag{4.61}$$

(For the definition of $\mathcal{P}_i$, see the comments following Lemma 4.5.) We have

$$\begin{aligned}\mathcal{P}_0\mathcal{P}_1\cdots\mathcal{P}_{l-1} &= \prod_{i=1}^{j-1}(\mathcal{P}_{r_{i-1}}\cdots\mathcal{P}_{r_i-2})(\mathcal{P}_{r_{j-1}}\cdots\mathcal{P}_{l-1}) \\ &= \left(\prod_{i=1}^{j-1} 2^{\sum_{z=1}^{r_i-r_{i-1}-1} 2^{n-r_i+z}}\right) 2^{\sum_{z=0}^{l-r_{j-1}-1} 2^{n-l+z}}.\end{aligned} \tag{4.62}$$

By equations (4.57)–(4.60) and (4.62) a straightforward algebraic simplification of the right hand side of equation (4.61) gives $\mathcal{N}_{k_{min}}(L, L_{j,l,C}) = 2^{\rho(j,l,C)}$ with $\rho(j,l,C)$ as in equation (4.54). We note that the condition in equation (4.60) is necessary to avoid double counting in determining the number of distinct possibilities for $\psi^l(\mathbf{S})$ over all $\mathbf{S} \in \mathcal{A}(L)$ such that $\psi^{l+1}(\mathbf{S})$ and $\psi^{l+1}(\mathbf{S}')$ are fixed.

Case 2: $w = t + 1$
In this case we note that the two possibilities for vectors in the $(n-1)$-th step of the Games-Chan algorithm are 01 and 10. Using this it can be shown that the expression for $\mathcal{D}^1(l)$ in equation (4.57) holds for $w = t + 1$. The remaining details are similar to those in Case 1.

To obtain $\mathcal{N}_{k_{min}}(L, 0)$ we only have to count the number of $\mathbf{S} \in \mathcal{A}(L)$ with $w_H(\mathbf{S}) = 2^t$. By equation (4.35) and property P1 the expression for $\mathcal{N}_{k_{min}}(L, 0)$ follows using an argument similar to the one used for finding $\mathcal{D}^1(l)$ as in equation (4.57).

This completes the proof of the theorem. □

4.2.3 Concluding Remarks

In this section we analyzed the Games-Chan algorithm and obtained a partial counting function for the number of 2^n-periodic binary sequences with a given linear complexity L and a given $k_{min}(L)$-error linear complexity. We believe that the full counting function can be obtained by using results in Section 4.1.3. Although we do not provide the counting function, we note that procedure in Theorem 4.18 also shows the existence of sequences with any given value of $k_{min}(L)$-error linear complexity derived in Theorem 4.16.

Recently Etzion et al. [16] obtained further results on the error linear complexity profiles of 2^n-periodic binary sequences using the costed binary sequences approach in the Lauder-Patterson [44] algorithm. The critical error linear complexity profile of a sequence $\mathbf{S}$ is the set of points $(k, L_k(\mathbf{S}))$ given by

$$\mathfrak{E}(\mathbf{S}) = \{(k, L_k(\mathbf{S})) : L_{k'}(\mathbf{S}) > L_k(\mathbf{S}) \quad \forall \quad k' < k\}.$$

Each point $(k, L_k(\mathbf{S})) \in \mathfrak{E}(\mathbf{S})$ is called a critical point. We see that $\mathfrak{E}(\mathbf{S})$ is the set of points where the linear complexity decreases. Etzion et al. obtained a formula for the number of sequences with exactly two critical points. They also showed $|\mathfrak{E}(\mathbf{S})| \leq 2^{n-2} + 2$ over all 2^n-periodic binary sequences $\mathbf{S}$. It is an open question whether this upper bound is attained for $n \geq 7$; computer experiments showed the

tightness for $n < 7$ [16]. It is also an open question to obtain the number of 2^n-periodic binary sequences with a given number of critical points. We believe further analysis of the Games-Chan algorithm using our approach might give insights into these questions. Generalizations of these results to sequences with arbitrary periods or other special periods would be helpful.

4.3 Sequences with Fixed 2-Error or 3-Error Linear Complexity

In contrast to the most common approach of analyzing efficient sequence complexity measure computation algorithms to obtain counting functions, Fu, Niederreiter, and Su [18] studied the linear complexity and the 1-error linear complexity of 2^n-periodic binary sequences to characterize such sequences with fixed 1-error linear complexity using algebraic and combinatorial methods. Su and Chen [81] used the same approach to obtain results for the 1-error linear complexity of p^n-periodic sequences over $\mathbb{F}_p$.

Fu et al. derived some properties of the set $\mathcal{A}(L)$ that deal with changing two symbols per period at fixed positions in sequences of $\mathcal{A}(L)$ and used them to obtain the characterization of $\mathcal{A}_1(L)$.

Theorem 4.19 ([18]). *Let* $\mathbf{E}_i$, $0 \le i \le 2^n - 1$, *be the* 2^n*-periodic binary sequence with a* 1 *at position* i *and* 0 *elsewhere in each period and* $\mathbf{0}$ *be the zero sequence. We have*

(i) $\mathcal{A}_1(0) = \{\mathbf{0}, \mathbf{E}_0, \ldots, \mathbf{E}_{2^n-1}\}$ *and* $\mathcal{N}_1(0) = 2^n + 1$.

(ii) If $2^n - 2^{n-r} < L < 2^n - 2^{n-r-1}$ *for some* $0 \le r \le n-2$, *then*

$$\mathcal{A}_1(L) = \mathcal{A}(L) \bigcup \left(\bigcup_{i=0}^{2^{n-r}-1} (\mathcal{A}(L) + \mathbf{E}_i) \right)$$

and $\mathcal{N}_1(L) = (2^{n-r} + 1)2^{L-1}$.

(iii) If $L = 2^n - 2^{n-r}$, $r = 1, 2, \ldots, n$, *then* $\mathcal{A}_1(L) = \mathcal{A}(L)$ *and* $\mathcal{N}_1(L) = 2^{L-1}$.

For the rest of this section we use the notation and auxiliary results from Section 4.1. In this section we first study the effect of t symbol changes in 2^n-periodic binary sequences for small t. Specifically, for various special cases of L we determine some t symbol changes of sequences in $\mathcal{A}(L)$ that result in sequences in $\mathcal{A}(L)$ for $t = 4$ and 6; the case when $t = 2$ is already handled in Section 4.1.3. We also characterize specific 2, 4, and 6 symbol changes in sequences of $\mathcal{A}(L)$ that result in 2^n-periodic binary sequences with linear complexity strictly less than L. We use these characterizations to construct disjoint decompositions of the sets $\mathcal{A}_2(L)$ and $\mathcal{A}_3(L)$ of sequences with fixed 2-error or 3-error linear complexity L. Each set in the decompositions arises by changing all sequences in $\mathcal{A}(L)$ in a fixed set of positions. Using the characterizations of $\mathcal{A}_2(L)$ and $\mathcal{A}_3(L)$ we determine the expressions for $\mathcal{N}_2(L)$ and $\mathcal{N}_3(L)$. A portion of this work was presented at the 5th international conference on *SEquences and Their Applications (SETA 2008)* [34] and full results appear in the journal *Designs, Codes, and Cryptography* [35].

4.3.1 Effect of Small Changes on the Linear Complexity

In this section we study the effect of a small number of changes on the linear complexity of sequences in $\mathcal{A}(L)$ and derive some properties of $\mathcal{A}(L)$, which extend those in Fu et al.'s paper [18]. We use the following generalization of Fu et al.'s result [18, Theorem 1] in later sections.

Theorem 4.20. *For a given $r \in \{1, \ldots, n-1\}$, let $1 \le L < 2^{n-r}$. Then for any two distinct sequences $\mathbf{S}, \mathbf{S}' \in \mathcal{A}(L)$ we have*

$$d_H(\mathbf{S}, \mathbf{S}') = t \cdot 2^{r+1} \quad \textit{for some} \quad t \in \{1, 2, 3, \ldots, 2^{n-r-1}\},$$

which implies $d_H(\mathbf{S}, \mathbf{S}') \ge 2^{r+1}$.

Proof. For any sequence $\mathbf{S} \in \mathcal{A}(L)$, consider the corresponding polynomial $\mathbf{S}(x) = (1+x)^{2^n-L}a(x)$, where $a(x) \in \mathbb{F}_2[x]$ such that $\deg(a(x)) \le L-1$ and $a(1) \ne 0$. Since $1 \le L < 2^{n-r}$, we have $2^n - L > 2^n - 2^{n-r}$. The generating function for $\mathbf{S}$ is given by

$$\frac{\mathbf{S}(x)}{1-x^{2^n}} = \frac{(1+x)^{(2^n-2^{n-r})+(2^{n-r}-L)}a(x)}{(1+x)^{2^n}} = \frac{(1+x)^{2^{n-r}-L}a(x)}{1-x^{2^{n-r}}},$$

which implies 2^{n-r} is a period of $\mathbf{S}$. Corresponding to any sequence $\mathbf{M} \in \mathcal{A}(L)$, let $\mathbf{M}^{(2^{n-r})}$ denote the 2^{n-r}-tuple $(m_0, m_1, \ldots, m_{2^{n-r}-1})$. Since $1 \le L < 2^{n-r}$, from Lemma 4.9 we know that $w_H(\mathbf{S}^{(2^{n-r})})$ and $w_H(\mathbf{S}'^{(2^{n-r})})$ are even. Hence the Hamming distance between $\mathbf{S}^{(2^{n-r})}$ and $\mathbf{S}'^{(2^{n-r})}$ is even. That is $d_H(\mathbf{S}^{(2^{n-r})}, \mathbf{S}'^{(2^{n-r})}) = 2t$ for some $t \in \{1, 2, 3, \ldots, 2^{n-r-1}\}$. Since 2^{n-r} is a period of $\mathbf{S}$ and $\mathbf{S}'$, we have $d_H(\mathbf{S}, \mathbf{S}') = 2^r \cdot d_H(\mathbf{S}^{(2^{n-r})}, \mathbf{S}'^{(2^{n-r})}) = t \cdot 2^{r+1}$. This completes the proof of the theorem. □

Let $\mathbf{S}$ be a 2^n-periodic binary sequence with $0 < L(\mathbf{S}) < 2^n$ and let m be an integer such that $1 \le m \le n-1$. If $\mathtt{minerr}(\mathbf{S}) = 2^{m+1}$, then by Lemma 4.1 the linear complexity of $\mathbf{S}$ can be uniquely expressed as

$$L(\mathbf{S}) = 2^n - \sum_{i=1}^{m+1} 2^{n-r_i}, \tag{4.63}$$

where $1 \le r_1 < \cdots < r_{m+1} \le n$. If $\mathtt{minerr}(\mathbf{S}) \ge 2^{m+1}$, then by equation (4.63) the linear complexity is bounded as

$$2^n - \left(\sum_{i=1}^{m-1} 2^{n-r_i} + 2^{n-r_m+1} \right) < L(\mathbf{S}) < 2^n - \sum_{i=1}^{m} 2^{n-r_i}, \tag{4.64}$$

for some $r_i \in \{1, \ldots, n\}$, $i = 1, \ldots, m$, satisfying $1 \le r_1 < \cdots < r_m$. Note that conversely, for any sequence $\mathbf{S}$ satisfying the inequality (4.64), we have $\mathtt{minerr}(\mathbf{S}) \ge 2^{m+1}$. We also note that the bounds in (4.64) are unique in the sense that the linear complexity of any 2^n-periodic sequence $\mathbf{S}$ with $\mathtt{minerr}(\mathbf{S}) \ge 2^{m+1}$ satisfies exactly one inequality of the particular form given in equation (4.64). Note that by equation (4.64) any L such that $w_H(2^n - L) \ge 3$ can be bounded as $2^n - (2^{n-r_1} + 2^{n-r_2}) < L < 2^n - (2^{n-r_1} + 2^{n-r_2-1})$ for some $1 \le r_1 \le r_2 < n$.

The first main result of this section deals with extending Lemma 4.8 to the case when four symbols per period are changed. Here we describe four symbol changes for sequences in $\mathcal{A}(L)$ such that the linear complexity of the modified sequences remains L. We assume that the four positions where the changes are made are distinct since the cases of four symbol changes when more than two positions are identical are covered by Lemma 4.8. We also present a corresponding result when six symbols are changed. The proof uses the Games-Chan algorithm and the corresponding notation in Section 4.1.2.

Theorem 4.21. *Let* $\mathbf{S} \in \mathcal{A}(L)$ *where*

$$2^n - (2^{n-r_1} + 2^{n-r_2}) < L < 2^n - (2^{n-r_1} + 2^{n-r_2-1}), \tag{4.65}$$

for some r_1 *and* r_2 *satisfying* $1 \le r_1 \le r_2 < n$.

(i) Consider any four integers i, j, k, *and* l *such that* $0 \le i < j < k < l \le 2^{n-r_1+1} - 1$. *Then* $L(\mathbf{S}_{i,j,k,l}) = L(\mathbf{S})$ *if and only if* i, j, k, *and* l *are in the form*

$$i = u + g_1 2^{n-r_2}, \quad j = u + g_2 2^{n-r_2}, \quad k = i + 2^{n-r_1}, \quad \text{and} \quad l = j + 2^{n-r_1}, \tag{4.66}$$

where $0 \le u \le 2^{n-r_2} - 1$ *and* $0 \le g_1 < g_2 \le 2^{r_2-r_1} - 1$.

(ii) There do not exist integers $i_1, \ldots, i_6$ *such that* $0 \le i_1 < \cdots < i_6 \le 2^{n-r_1+1} - 1$ *and* $L(\mathbf{S}_{i_1,\ldots,i_6}) = L(\mathbf{S})$.

Proof. We only prove the forward direction of part (i) of the theorem. The other direction is straightforward and can be proved by reversing the argument used for the forward case.

Consider any sequence

$$\mathbf{S}_{i,j,k,l} \in \mathcal{A}(L), \quad \text{where} \quad 0 \le i < j < k < l \le 2^{n-r_1+1} - 1. \tag{4.67}$$

From equation (4.65) we have

$$w_H(2^n - L) \ge 3 \quad \text{and} \quad L = 2^n - (2^{n-r_1} + 2^{n-r_2-1} + c), \tag{4.68}$$

for some $0 < c < 2^{n-r_2-1}$. From equations (4.7), (4.8), and (4.68), we have

$$\forall\ \mathbf{S} \in \mathcal{A}(L), \quad \psi_L^{r_1-1}(\mathbf{S}) = \psi_R^{r_1-1}(\mathbf{S}) \quad \text{and} \quad \psi_L^{r_2}(\mathbf{S}) = \psi_R^{r_2}(\mathbf{S}). \tag{4.69}$$

By Lemma 4.5 and equation (4.68) the left and right halves are not equal during the first $r_1 - 2$ steps of the Games-Chan procedure for any $\mathbf{S} \in \mathcal{A}(L)$. Thus, since $0 \le i, j, k, l \le 2^{n-r_1+1} - 1$, by the procedure of the Games-Chan algorithm we get

$$d_H(\psi^{r_1-1}(\mathbf{S}), \psi^{r_1-1}(\mathbf{S}_{i,j,k,l})) = 4. \tag{4.70}$$

By equations (4.69) and (4.70), the four positions where the vectors $\psi^{r_1-1}(\mathbf{S})$ and $\psi^{r_1-1}(\mathbf{S}_{i,j,k,l})$ differ are of the form

$$c_1, \quad c_2, \quad c_1 + 2^{n-r_1}, \quad \text{and} \quad c_2 + 2^{n-r_1}, \quad \text{for some} \quad 0 \le c_1 < c_2 \le 2^{n-r_1} - 1. \tag{4.71}$$

From equations (4.69) and (4.70), we have $d_H(\psi_L^{r_1-1}(\mathbf{S}), \psi_L^{r_1-1}(\mathbf{S}_{i,j,k,l})) = 2$. This implies

$$d_H(\psi^{r_1}(\mathbf{S}), \psi^{r_1}(\mathbf{S}_{i,j,k,l})) = 2. \tag{4.72}$$

Now we treat $\psi^{r_1}(\mathbf{S})$ and $\psi^{r_1}(\mathbf{S}_{i,j,k,l})$ as the first periods of 2^{n-r_1}-periodic binary sequences $\mathbf{S}'$ and $\mathbf{S}'_{i,j,k,l}$, respectively, that differ at 2 positions. With this notation, from equations (4.71) and (4.72) we have $\mathbf{S}' = (\psi^{r_1}(\mathbf{S}))^\infty$, $\mathbf{S}'_{i,j,k,l} = (\psi^{r_1}(\mathbf{S}_{i,j,k,l}))^\infty$, and

$$\mathbf{S}'_{i,j,k,l}(x) = \mathbf{S}'(x) + x^{c_1} + x^{c_2}. \tag{4.73}$$

As a consequence of the procedure of the Games-Chan algorithm, since the left and right halves are different in the first $r_1 - 2$ steps for both $\mathbf{S}$ and $\mathbf{S}_{i,j,k,l}$, we have

$$m^{r_1-1}(\mathbf{S}) = m^{r_1-1}(\mathbf{S}_{i,j,k,l}) = 2^{n-1} + \cdots + 2^{n-r_1+1} = 2^n - 2^{n-r_1+1}. \tag{4.74}$$

Using Lemma 4.5 and by equations (4.67), (4.73), and (4.74) we have

$$\mathbf{S}', \mathbf{S}'_{i,j,k,l} \in \mathcal{A}(L') \quad \text{where} \quad L' = L - (2^n - 2^{n-r_1+1}). \tag{4.75}$$

Equations (4.65) and (4.75) imply that L' satisfies

$$2^{n-r_1} - 2^{n-r_2} < L' < 2^{n-r_1} - 2^{n-r_2-1}. \tag{4.76}$$

By Lemma 4.8 and equation (4.76), the positions c_1 and c_2 in equations (4.71) and (4.73) must be in the form

$$\begin{aligned} c_i &= u + g_i 2^{n-r_2}, \quad i = 1, 2, \quad \text{where} \\ 0 \le u &\le 2^{n-r_2} - 1 \quad \text{and} \quad 0 \le g_1 < g_2 \le 2^{r_2-r_1} - 1. \end{aligned} \tag{4.77}$$

From equations (4.71) and (4.77), the four positions, denoted f_1, f_2, f_3, and f_4, where $\psi^{r_1-1}(\mathbf{S})$ and $\psi^{r_1-1}(\mathbf{S}_{i,j,k,l})$ differ are of the form

$$f_1 = c_1, \quad f_2 = c_2, \quad f_3 = c_1 + 2^{n-r_1}, \quad \text{and} \quad f_4 = c_2 + 2^{n-r_1}, \tag{4.78}$$

where c_1 and c_2 are as in equation (4.77).

From the procedure of the Games-Chan algorithm, we observe that a symbol change at any position c in $\psi^{r_1-1}(\mathbf{S})$, $0 \le c \le 2^{n-r_1+1} - 1$, can be effected by changing the symbol at one of the corresponding positions $(c + b2^{n-r_1+1}) \bmod 2^n$, $b \in \{0, \ldots, 2^{r_1-1} - 1\}$, in each period of $\mathbf{S}$. Thus from equations (4.77) and (4.78), i, j, k, and l must be in the form given in equation (4.66).

To prove part (ii) assume that there exist integers $i_1, \ldots, i_6$ such that $0 \le i_1 < \cdots < i_6 \le 2^{n-r_1+1} - 1$ and

$$L(\mathbf{S}_{i_1,\ldots,i_6}) = L(\mathbf{S}). \tag{4.79}$$

From the procedure of the Games-Chan algorithm, using an argument similar to that used to arrive at equation (4.72), we have

$$d_H(\psi^{r_1}(\mathbf{S}), \psi^{r_1}(\mathbf{S}_{i_1,\ldots,i_6})) = 3. \tag{4.80}$$

By equation (4.68) and Lemma 4.5 we know $w_H(\psi^{r_1}(\mathbf{S}))$ is even since otherwise $L(\mathbf{S}) = 2^n - 2^{n-r_1}$. Using this, equation (4.80) implies that $w_H(\psi^{r_1}(\mathbf{S}_{i_1,\ldots,i_6}))$ is odd, which contradicts equation (4.79). Thus part (ii) of the theorem is proved. □

Remark 4.3. Note that when $r_1 = r_2 = 1$ in Theorem 4.21(i), there are no possible distinct values for g_1 and g_2 in equation (4.66). Thus when $0 < L < 2^{n-2}$ there do not exist distinct four symbol changes to any sequence in $\mathcal{A}(L)$ that result in sequences with linear complexity L. This is an alternative proof of Theorem 4.20 when $r = 2$.

Also, for some values of L in equation (4.65), in order to write L in the form as in equation (4.64), we must allow $r_1 = r_2$.

The following result describes certain four and six symbol changes for sequences in $\mathcal{A}(L)$ that retain the linear complexity. The proof is omitted as it can be proved using Lemma 4.7 and the approach used in Theorem 4.21.

Theorem 4.22. *Let $\mathbf{S} \in \mathcal{A}(L)$ where $L = 2^n - (2^{n-r_1} + 2^{n-r_2})$ for some r_1 and r_2 such that $1 \le r_1 < r_2 \le n$.*

(i) Consider any four integers i, j, k, and l such that $0 \le i < j < k < l \le 2^{n-r_1+1} - 1$. Then $L(\mathbf{S}_{i,j,k,l}) = L(\mathbf{S})$ if and only if i, j, k, and l are in the form

$$\begin{aligned} i = u + g_1 2^{n-r_2+1}, \quad j = u + g_2 2^{n-r_2+1}, \\ k = i + 2^{n-r_1}, \quad \text{and} \quad l = j + 2^{n-r_1}, \end{aligned} \tag{4.81}$$

where

$$0 \le u \le 2^{n-r_2+1} - 1 \quad \text{and} \quad 1 \le g_1 < g_2 \le 2^{r_2-r_1-1} - 1. \tag{4.82}$$

(ii) There do not exist integers $i_1, \ldots, i_6$ such that $0 \le i_1 < \cdots < i_6 \le 2^{n-r_1+1} - 1$ and $L(\mathbf{S}_{i_1,\ldots,i_6}) = L(\mathbf{S})$.

For any polynomial $a(x) \in \mathbb{F}_2[x]$ given by $a(x) = 1 + x^{a_1} + \cdots + x^{a_{q-1}}$, define the weight $W(a(x)) = q$. Next we handle two symbol changes that decrease the linear complexity of 2^n-periodic binary sequences.

Lemma 4.23. *For any sequence $\mathbf{S} \in \mathcal{A}(L)$, where $L = 2^n - 2^{n-r}$ for some $1 \le r \le n$, and for any integer $0 \le i \le 2^n - 1$, the number of sequences $\mathbf{S}_{i,j}$ such that $L(\mathbf{S}_{i,j}) < L$, where $0 \le j \le 2^n - 1$ and $j \ne i$, is exactly 2^{r-1}, corresponding to all $j \in \{(i + (2t+1)2^{n-r}) \bmod 2^n : 0 \le t \le 2^{r-1} - 1\}$.*

Proof. First we prove the forward direction of the result. Let $\mathbf{S}(x) = (1+x)^{2^{n-r}} a(x)$ for some $a(x) \in \mathbb{F}_2[x]$ such that $\deg(a(x)) \le 2^n - 2^{n-r} - 1$ and $a(1) = 1$. The corresponding polynomial for $\mathbf{S}_{i,j}$ is

$$\mathbf{S}_{i,j}(x) = (1+x)^{2^{n-r}} a(x) + x^i + x^j.$$

So $L(\mathbf{S}_{i,j}) = 2^n - \deg(\gcd((1+x)^{2^n}, (1+x)^{2^{n-r}} a(x) + x^i + x^j))$ and hence we have

$$L(\mathbf{S}_{i,j}) < L \quad \text{if and only if} \quad \gcd((1+x)^{2^n}, x^i + x^j) = (1+x)^{2^{n-r}}. \tag{4.83}$$

Without loss of generality we may assume $i < j$. It is a well known fact that $\gcd(1 + x^a, 1 + x^b) = 1 + x^{\gcd(a,b)}$. Hence we get

$$\gcd((1+x)^{2^n}, x^i + x^j) = \gcd(1 + x^{2^n}, 1 + x^{j-i}) = 1 + x^{\gcd(2^n, j-i)} = 1 + x^{2^{n-r}}$$

if and only if 2^{n-r} divides $j-i$ and no higher power of 2 divides $j-i$. Thus equation (4.83) implies that $L(\mathbf{S}_{i,j}) < L$ if and only if $j = i + d2^{n-r}$ for some odd integer d, which proves the forward direction. The reverse direction can be proved using an argument similar to that used in proving the reverse direction of Lemma 4.7. □

Corollary 4.24. *For any sequence* $\mathbf{S} \in \mathcal{A}(L)$*, where* $L = 2^n - 2^{n-r}$ *for some* $1 \le r \le n$*, there are* 2^{n+r-2} *distinct pairs* i*,* j*,* $0 \le i < j \le 2^n - 1$*, such that* $L(\mathbf{S}_{i,j}) < L$*. All such* i*,* j *are described as*

$$i \quad \text{and} \quad j = i + (2t+1)2^{n-r}, \tag{4.84}$$

where

$$0 \le i \le 2^n - 2^{n-r} - 1 \quad \text{and} \quad 0 \le t \le 2^{r-1} - 1 - \lceil (\lfloor i/2^{n-r} \rfloor)/2 \rceil. \tag{4.85}$$

Also, the distinct pairs i*,* j*,* $0 \le i < j \le 2^n - 1$*, such that*

$$1 + x^{j-i} = (1+x)^{2^{n-r}} b(x) \tag{4.86}$$

for some $b(x) \in \mathbb{F}_2[x]$ *with* $b(1) = 1$ *and* $\deg(b(x)) \le 2^n - 2^{n-r} - 1$*, are exactly those described in equations* (4.84) *and* (4.85)*.*

Proof. By Lemma 4.23 for each $i \ge 2^n - 2^{n-r}$ there is no j such that $i < j \le 2^n - 1$ and $L(\mathbf{S}_{i,j}) < L$. Also, for each $0 \le i \le 2^n - 2^{n-r} - 1$ there are exactly $2^{r-1} - \lceil (\lfloor i/2^{n-r} \rfloor)/2 \rceil$ odd multiples of 2^{n-r} corresponding to $0 \le t \le 2^{r-1} - 1 - \lceil (\lfloor i/2^{n-r} \rfloor)/2 \rceil$ such that $L(\mathbf{S}_{i,i+(2t+1)2^{n-r}}) < L$. Thus all i, j, $0 \le i < j \le 2^n - 1$, such that $L(\mathbf{S}_{i,j}) < L$ are as described in equations (4.84) and (4.85).

The number of distinct pairs i, j obtained from equations (4.84) and (4.85) is

$$\begin{aligned} \sum_{i=0}^{2^n-1} \left(2^{r-1} - \lceil (\lfloor i/2^{n-r} \rfloor)/2 \rceil\right) &= \sum_{i=0}^{2^{n-r}-1} 2^{r-1} + \sum_{l=1}^{2^{r-1}-1} \left(\sum_{i=(2l-1)2^{n-r}}^{(2l+1)2^{n-r}-1} (2^{r-1} - l) \right) \\ &= 2^{n-r} 2^{r-1} + 2^{n-r+1} \left(\sum_{l=1}^{2^{r-1}-1} (2^{r-1} - l) \right) \\ &= 2^{n+r-2}. \end{aligned} \tag{4.87}$$

By the definition of linear complexity it is straightforward to see that the integers i, j in equations (4.84) and (4.85) are exactly those that satisfy equation (4.86). □

Our next result deals with four symbol changes that decrease the linear complexity of 2^n-periodic binary sequences.

Theorem 4.25. *Let* $\mathbf{S} \in \mathcal{A}(L)$ *where* $L = 2^n - (2^{n-r_1} + 2^{n-r_2})$ *for some* r_1*,* r_2 *such that* $1 \le r_1 < r_2 \le n$*.*

(i) Consider any four integers i*,* j*,* k*, and* l *such that* $0 \le i < j < k < l \le 2^{n-r_1+1} - 1$*. Then* $L(\mathbf{S}_{i,j,k,l}) < L$ *if and only if* i*,* j*,* k*, and* l *are in the form*

$$i, \quad j = i + (2t+1)2^{n-r_2}, \quad k = i + 2^{n-r_1}, \quad \text{and} \quad l = j + 2^{n-r_1}, \tag{4.88}$$

where

$$0 \le i \le 2^{n-r_1} - 2^{n-r_2} - 1 \quad \textit{and} \quad 0 \le t \le 2^{r_2-r_1-1} - 1 - \lceil (\lfloor i/2^{n-r_2} \rfloor)/2 \rceil. \tag{4.89}$$

Furthermore, if $\mathcal{K}(L)$ is the set of four symbol changes to $\mathbf{S}$ described in equations (4.88) and (4.89) that decrease its linear complexity, then

$$\begin{aligned} &|\mathcal{K}(L)| \\ &= |\{\{i,j,k,l\} : 0 \le i < j < k < l < 2^{n-r_1+1} \quad \textit{and} \quad L(\mathbf{S}_{i,j,k,l}) < L\}| \\ &= 2^{n+r_2-2r_1-2}. \end{aligned} \tag{4.90}$$

(ii) For any four integers i_t, $t = 1, \ldots, 4$, such that $0 \le i_1 < i_2 < i_3 < i_4 \le 2^n - 1$, we have $L(\mathbf{S}_{i_1,i_2,i_3,i_4}) < L$ if and only if $\{i_t \bmod 2^{n-r_1+1} : t = 1, \ldots, 4\} \in \mathcal{K}(L)$.

(iii) There do not exist integers $i_1, \ldots, i_6$, $0 \le i_1 < \cdots < i_6 \le 2^n - 1$, such that $L(\mathbf{S}_{i_1,\ldots,i_6}) < L$.

Proof. First we prove the forward direction of part (i). Consider the polynomial $\mathbf{S}(x) = (1+x)^{2^{n-r_1}+2^{n-r_2}} a(x)$ for some $a(x) \in \mathbb{F}_2[x]$ such that $\deg(a(x)) \le 2^n - 2^{n-r_1} - 2^{n-r_2} - 1$ and $a(1) = 1$. The corresponding polynomial for $\mathbf{S}_{i,j,k,l}$ is

$$\mathbf{S}_{i,j,k,l}(x) = (1+x)^{2^{n-r_1}+2^{n-r_2}} a(x) + x^i + x^j + x^k + x^l.$$

So $L(\mathbf{S}_{i,j,k,l}) = 2^n - \gcd((1+x)^{2^n}, (1+x)^{2^{n-r_1}+2^{n-r_2}} a(x) + x^i + x^j + x^k + x^l)$ and hence $L(\mathbf{S}_{i,j,k,l}) < L$ if and only if $\gcd((1+x)^{2^n}, x^i + x^j + x^k + x^l) = (1+x)^{2^{n-r_1}+2^{n-r_2}}$. This holds if and only if

$$\begin{aligned} 1 + x^{j-i} + x^{k-i} + x^{l-i} &= (1+x)^{2^{n-r_1}+2^{n-r_2}} b(x) \\ &= (1 + x^{2^{n-r_2}}) b(x) + x^{2^{n-r_1}} (1 + x^{2^{n-r_2}}) b(x) \end{aligned} \tag{4.91}$$

for some $b(x) \in \mathbb{F}_2[x]$ such that $b(1) = b(0) = 1$. Since $0 \le i < j < k < l \le 2^{n-r_1+1} - 1$ we have

$$\deg(b(x)) \le 2^{n-r_1} - 2^{n-r_2} - 1. \tag{4.92}$$

Since $W((1 + x^{2^{n-r_2}}) b(x)) \ge 2$, by equations (4.91) and (4.92) we see that

$$1 + x^{j-i} = (1 + x^{2^{n-r_2}}) b(x). \tag{4.93}$$

By Corollary 4.24 and equations (4.91), (4.92), and (4.93) we see that i, j, k, and l should be as in equation (4.88). The proof of the reverse direction of part (i) is straightforward and is similar to the proof of the reverse direction of Lemma 4.7. Equation (4.90) follows from equations (4.88), (4.89), (4.92), (4.93), Lemma 4.23, and an argument similar to that used in Corollary 4.24 by replacing n by $n - r_1$ and r by $r_2 - r_1$ in equation (4.87).

To prove the forward direction of part (ii), we first note that $L(\mathbf{S}_{i_1,i_2,i_3,i_4}) < L$ if and only if the polynomial

$$e(x) = x^{i_1} + x^{i_2} + x^{i_3} + x^{i_4} = (1+x)^{2^{n-r_1}+2^{n-r_2}} b'(x) \tag{4.94}$$

for some $b'(x) \in \mathbb{F}_2[x]$ such that $\deg(b'(x)) \leq 2^n - 2^{n-r_1} - 2^{n-r_2} - 1$ and $b'(1) = 1$. Let u be the largest power of $(1+x)$ dividing

$$e'(x) = x^{i_1 \bmod 2^{n-r_1+1}} + x^{i_2 \bmod 2^{n-r_1+1}} + x^{i_3 \bmod 2^{n-r_1+1}} + x^{i_4 \bmod 2^{n-r_1+1}} \tag{4.95}$$

so that

$$e'(x) = (1+x)^u b''(x) \tag{4.96}$$

for some $b''(x) \in \mathbb{F}_2[x]$ such that $\deg(b(x)) \leq 2^{n-r_1+1} - u$ and $b''(1) = 1$. For $t = 1, \ldots, 4$ denoting $q_t = \lfloor i_t / 2^{n-r_1+1} \rfloor$ we have

$$\begin{aligned} x^{i_t \bmod 2^{n-r_1+1}} + x^{i_t} &= x^{i_t \bmod 2^{n-r_1+1}} + x^{i_t \bmod 2^{n-r_1+1} + q_t 2^{n-r_1+1}} \\ &= x^{i_t \bmod 2^{n-r_1+1}} (1+x)^{2^{n-r_1+1}} (1 + \cdots + x^{q_t - 1})^{2^{n-r_1+1}}. \end{aligned}$$

By equations (4.94) and (4.96), this implies

$$(1+x)^{2^{n-r_1+1}} \mid e(x) + e'(x).$$

So

$$u = 2^{n-r_1} + 2^{n-r_2} \tag{4.97}$$

since $2^{n-r_1+1} > 2^{n-r_1} + 2^{n-r_2}$. Since $L = 2^n - (2^{n-r_1} + 2^{n-r_2})$, by equations (4.95)–(4.97), and the definition of linear complexity we see that the four symbol changes at positions $i_t \mod 2^{n-r_1+1}$, $t = 1, \ldots, 4$, lower the linear complexity of any $\mathbf{S} \in \mathcal{A}(L)$. Thus $\{i_t \bmod 2^{n-r_1+1} : t = 1, \ldots, 4\} \in \mathcal{K}(L)$, which concludes the proof of the forward direction of part (ii). The reverse direction of part (ii) can be proved similarly.

To prove part (iii), suppose there were integers $i_1, \ldots, i_6$, $0 \leq i_1 < \cdots < i_6 \leq 2^{n-r_1+1} - 1$, such that $L(\mathbf{S}_{i_1,\ldots,i_6}) < L$. By the argument used to arrive at equation (4.91) we have

$$x^{i_1} + \cdots + x^{i_6} = (1 + x^{2^{n-r_2}})c(x) + x^{2^{n-r_1}}(1 + x^{2^{n-r_2}})c(x), \tag{4.98}$$

for some $c(x) \in \mathbb{F}_2[x]$ such that $c(1) = 1$ and $\deg(c(x)) \leq 2^{n-r_1} - 2^{n-r_2} - 1$. By equation (4.98) and the upper bound on $\deg(c(x))$ it follows that $(1 + x^{2^{n-r_2}})c(x) = x^{i_1} + x^{i_2} + x^{i_3}$, which is not possible since $(1 + x^{2^{n-r_2}})c(x)$ has an even number of terms. So the result follows when $0 \leq i_1 < \cdots < i_6 \leq 2^{n-r_1+1} - 1$. The result holds even when $0 \leq i_1 < \cdots < i_6 \leq 2^n - 1$ due to an argument similar to that used to prove part (ii). □

Remark 4.4. Theorem 4.21 can also be proved with the approach of Theorem 4.25 by using results on polynomial weights [43, Proposition 3.2].

4.3.2 Additional Notation and Auxiliary Results

In this section we establish additional notation used for the rest of the section and derive some auxiliary results on the k-error linear complexity of 2^n-periodic binary sequences.

Recall that $\mathcal{A}_k(L)$ is the set of 2^n-periodic binary sequences with k-error linear complexity L and $\mathcal{N}_k(L) = |\mathcal{A}_k(L)|$. For any $1 \le t \le 2^n$, let $\mathbf{E}_{i_1,\ldots,i_t}$, $0 \le i_1 < \cdots < i_t \le 2^n - 1$, denote the 2^n-periodic binary sequence of weight t with a 1 at positions with subscripts $i_1, \ldots, i_t$ in the first period and 0 elsewhere. Further let $\mathbb{E}_t = \{\mathbf{E}_{i_1,\ldots,i_t} : 0 \le i_1 < i_2 < \cdots < i_t \le 2^n - 1\}$ for $t \ge 1$ and $\mathbb{E}_0 = \{\mathbf{0}\}$. We denote by $\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_t}$ the set $\{\mathbf{S} + \mathbf{E}_{i_1,\ldots,i_t} : \mathbf{S} \in \mathcal{A}(L)\}$. For the rest of this chapter, for any set $\mathcal{R}$ of 2^n-periodic binary sequences, by $\mathcal{A}(L)[\mathcal{R}]$ denote the set of sets $\{\mathcal{A}(L) + \mathbf{R} : \mathbf{R} \in \mathcal{R}\}$.

We have a straightforward result that will be used in the next few sections.

Lemma 4.26 ([53]). *For any 2^n-periodic binary sequence $\mathbf{S}$ and for $k \ge 2$, $L_k(\mathbf{S})$ is different from $2^n - 2^t$ for every integer t with $0 \le t < n$.*

We derive two auxiliary results used for the main results in the rest of this section. First we see that for fixed L, the sets $\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_t}$ form a partition of the set of sequences with period 2^n.

Theorem 4.27. *Let $\{i_1, \ldots, i_{t_1}\}$ and $\{j_1, \ldots, j_{t_2}\}$ denote two sets of subscripts where $0 \le i_l, j_m \le 2^n - 1$ for $l = 1, \ldots, t_1$ and $m = 1, \ldots, t_2$. Then*

$$(\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_{t_1}}) \cap (\mathcal{A}(L) + \mathbf{E}_{j_1,\ldots,j_{t_2}}) = \emptyset$$

or

$$\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_{t_1}} = \mathcal{A}(L) + \mathbf{E}_{j_1,\ldots,j_{t_2}}.$$

Proof. We assume

$$0 < L \le 2^n \tag{4.99}$$

since the result holds trivially for $L = 0$.

Suppose $(\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_{t_1}}) \cap (\mathcal{A}(L) + \mathbf{E}_{j_1,\ldots,j_{t_2}}) \ne \emptyset$. So there exist sequences $\mathbf{S}$, $\mathbf{S}' \in \mathcal{A}(L)$ such that $\mathbf{S} + \mathbf{E}_{i_1,\ldots,i_{t_1}} = \mathbf{S}' + \mathbf{E}_{j_1,\ldots,j_{t_2}}$. This implies that

$$\mathbf{S} + \mathbf{E}_{i_1,\ldots,i_{t_1}} + \mathbf{E}_{j_1,\ldots,j_{t_2}} = \mathbf{S}'. \tag{4.100}$$

Consider the corresponding polynomials of $\mathbf{S}$ and $\mathbf{S}'$ given by

$$\mathbf{S}(x) = (1-x)^{2^n - L} a(x) \quad \text{and} \quad \mathbf{S}'(x) = (1-x)^{2^n - L} a'(x), \tag{4.101}$$

where $a(1) = a'(1) = 1$. From equations (4.99) and (4.101) we have

$$\deg(\gcd((1-x^{2^n}), \mathbf{S}(x) + \mathbf{S}'(x))) > 2^n - L. \tag{4.102}$$

From equations (4.100) and (4.102) we have

$$\deg(\gcd((1-x^{2^n}), x^{i_1} + \cdots + x^{i_{t_1}} + x^{j_1} + \cdots + x^{j_{t_2}})) > 2^n - L. \tag{4.103}$$

To prove the theorem we first show that every sequence in $\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_{t_1}}$ is in $\mathcal{A}(L) + \mathbf{E}_{j_1,\ldots,j_{t_2}}$. Consider any $\mathbf{R} \in \mathcal{A}(L)$ with the corresponding polynomial

$$\mathbf{R}(x) = (1-x)^{2^n - L} b(x), \quad \text{where} \quad b(1) = 1. \tag{4.104}$$

Then let $\mathbf{R}' = \mathbf{R} + \mathbf{E}_{i_1,\ldots,i_{t_1}} + \mathbf{E}_{j_1,\ldots,j_{t_2}}$ with the corresponding polynomial $\mathbf{R}'(x)$. By equations (4.103) and (4.104) we have

$$\begin{aligned} &\deg(\gcd((1-x^{2^n}), \mathbf{R}'(x))) \\ &= \deg(\gcd((1-x)^{2^n}, \mathbf{R}(x) + x^{i_1} + \cdots + x^{i_{t_1}} + x^{j_1} + \cdots + x^{j_{t_2}})) \\ &= 2^n - L. \end{aligned} \tag{4.105}$$

From equation (4.105), using the definition of linear complexity we have $\mathbf{R}' \in \mathcal{A}(L)$, which implies $\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_{t_1}} \subseteq \mathcal{A}(L) + \mathbf{E}_{j_1,\ldots,j_{t_2}}$. By symmetry $\mathcal{A}(L) + \mathbf{E}_{j_1,\ldots,j_{t_2}} \subseteq \mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_{t_1}}$, which proves the theorem. □

We need the following generalization of Theorem 4 in Fu et al.'s paper [18] in later sections.

Lemma 4.28. *Let* $\mathbf{S}$ *be a* T*-periodic binary sequence. Consider any two positive integers* u*,* v *such that* $0 \le v \le u$ *and* $u + v < \mathtt{minerr}(\mathbf{S})$*. Then for any* T*-periodic binary sequence* $\mathbf{E}$ *such that* $w_H(\mathbf{E}) = v$ *we have*

$$L_u(\mathbf{S} + \mathbf{E}) = L(\mathbf{S}).$$

Proof. First we note that $L_i(\mathbf{S}) = L(\mathbf{S})$ for $i = 0, \ldots, \mathtt{minerr}(\mathbf{S}) - 1$. Since $u + v < \mathtt{minerr}(\mathbf{S})$, by definitions of $L_u(\mathbf{S})$ and $L_{u+v}(\mathbf{S})$ we get

$$L_u(\mathbf{S} + \mathbf{E}) \ge L_{u+v}(\mathbf{S}) = L(\mathbf{S}). \tag{4.106}$$

Also, from the observations that $(\mathbf{S} + \mathbf{E}) + \mathbf{E} = \mathbf{S}$ and $w_H(\mathbf{E}) = v \le u$, we get

$$L_u(\mathbf{S} + \mathbf{E}) \le L(\mathbf{S}). \tag{4.107}$$

The lemma follows from equations (4.106) and (4.107). □

Next we prove a result on the characterization and counting function of $\mathcal{A}_k(L)$ for certain specific values of k and L.

Theorem 4.29. *Consider* $L \ge 0$ *such that* $w_H(2^n - L) \ge r+1$ *for some* $0 \le r \le n-1$*.*

(i) The set

$$\mathcal{A}_k(L) = \bigcup_{t=0}^{k} \left(\bigcup_{\mathbf{E}_{i_1,\ldots,i_t} \in \mathbb{E}_t} (\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_t}) \right) \quad \textit{for} \quad k = 1, \ldots, 2^r - 1. \tag{4.108}$$

(ii) Furthermore, if $1 \le L < 2^{n-r}$ *then the sets* $\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_t}$*,* $\mathbf{E}_{i_1,\ldots,i_t} \in \mathbb{E}_t$ *for* $t = 0, \ldots, 2^r - 1$ *are disjoint and*

$$\mathcal{N}_k(L) = \left(\sum_{i=0}^{k} \binom{2^n}{i} \right) 2^{L-1} \quad \textit{for} \quad k = 1, \ldots, 2^r - 1. \tag{4.109}$$

Proof. By Lemma 4.1 and the hypothesis $w_H(2^n - L) \geq r + 1$ for any $\mathbf{S} \in \mathcal{A}(L)$ we have $\mathtt{minerr}(\mathbf{S}) \geq 2^{r+1}$. Using this and Lemma 4.28 we have

$$\bigcup_{t=0}^{k} \left(\bigcup_{\mathbf{E}_{i_1,\ldots,i_t} \in \mathbb{E}_t} (\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_t}) \right) \subseteq \mathcal{A}_k(L).$$

Using this, equation (4.108) follows from the definition of k-error linear complexity.

To show part (ii) assume that $1 \leq L < 2^{n-r}$. To show that the sets $\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_t}$, $\mathbf{E}_{i_1,\ldots,i_t} \in \mathbb{E}_t$, $t = 0, \ldots, 2^r - 1$, are all disjoint, by Theorem 4.27, it is enough to show that no two of these sets are equal. We show this by contradiction. Any two sets $\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_u}$ and $\mathcal{A}(L) + \mathbf{E}_{j_1,\ldots,j_v}$, $0 \leq u, v \leq 2^r - 1$, are equal if and only if

$$\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_u,j_1,\ldots,j_v} = \mathcal{A}(L) \quad \text{with} \quad u + v \leq 2^{r+1} - 2. \tag{4.110}$$

By Theorem 4.20 for any two sequences $\mathbf{S}, \mathbf{S}' \in \mathcal{A}(L)$ we have $d_H(\mathbf{S}, \mathbf{S}') \geq 2^{r+1}$. Thus the set equality in equation (4.110) does not hold and all the sets $\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_t}$, $\mathbf{E}_{i_1,\ldots,i_t} \in \mathbb{E}_t$, $t = 0, \ldots, 2^r - 1$, are disjoint. Using this, the counting function in equation (4.109) follows from equation (4.108). □

4.3.3 Characterization When $w_H(2^n - L) \neq 2$

Here we characterize the 2^n-periodic binary sequences with fixed 2-error or 3-error linear complexity when the linear complexity is not of the form $2^n - (2^i + 2^j)$, $0 \leq i < j \leq n - 1$, by using the results from the previous subsection. First we obtain the results for 2-error linear complexity and then we extend them to the 3-error linear complexity case.

It is straightforward to see that

$$\mathcal{A}_2(0) = \mathbb{E}_1 \cup \mathbb{E}_2 \cup \{\mathbf{0}\} \quad \text{and} \quad \mathcal{N}_2(0) = \binom{2^n}{2} + 2^n + 1.$$

From Lemmas 4.9 and 4.11 we have $\mathcal{A}_2(2^n) = \emptyset$ and $\mathcal{N}_2(2^n) = 0$. From Lemma 4.26 we get $\mathcal{A}_2(L) = \emptyset$ and $\mathcal{N}_2(L) = 0$ for $L = 2^n - 2^t$, $0 \leq t < n$. Thus it remains to consider the case when $w_H(2^n - L) \geq 3$.

For any $1 \leq L < 2^{n-1}$, from Theorem 4.20 we know that for any two sequences $\mathbf{S}, \mathbf{S}' \in \mathcal{A}(L)$, $d_H(\mathbf{S}, \mathbf{S}') \geq 4$. Hence we have

$$\mathcal{A}(L) \cap (\mathcal{A}(L) + \mathbf{E}_t) = \emptyset, \tag{4.111}$$

$$\mathcal{A}(L) \cap (\mathcal{A}(L) + \mathbf{E}_{i,j}) = \emptyset, \quad \text{and} \tag{4.112}$$

$$(\mathcal{A}(L) + \mathbf{E}_t) \cap (\mathcal{A}(L) + \mathbf{E}_{i,j}) = \emptyset, \tag{4.113}$$

for all $\mathbf{E}_t \in \mathbb{E}_1$ and $\mathbf{E}_{i,j} \in \mathbb{E}_2$.

Theorem 4.30. *Let $w_H(2^n - L) \geq 3$ where*

$$2^n - (2^{n-r_1} + 2^{n-r_2}) < L < 2^n - (2^{n-r_1} + 2^{n-r_2-1}), \tag{4.114}$$

for some r_1 and r_2 satisfying $1 \le r_1 \le r_2 \le n-1$. Then

$$\mathcal{A}_2(L) = \mathcal{A}(L) \bigcup \left(\bigcup_{\mathbf{E}_i \in \mathbb{E}_1} (\mathcal{A}(L) + \mathbf{E}_i) \right) \bigcup \left(\bigcup_{\mathbf{E}_{i,j} \in \mathbb{E}_2} (\mathcal{A}(L) + \mathbf{E}_{i,j}) \right). \tag{4.115}$$

Define the sets

$$\begin{aligned} \mathbb{D}_1(L) &= \{\mathbf{E}_i : 0 \le i \le 2^{n-r_1+1} - 1\} \quad \textit{and} \\ \mathbb{D}_2(L) &= \{\mathbf{E}_{i,j} : 0 \le i < j \le 2^{n-r_1+1} - 1\}, \end{aligned} \tag{4.116}$$

where the definitions implicitly depend on L. Define the sets $\mathcal{D}^1(L)$ and $\mathcal{D}^2(L)$ by

$$\begin{aligned} \mathcal{D}^1(L) = \{&\mathbf{E}_{i,i+2^{n-r_1}} : i = u + t2^{n-r_2}, 1 \le t \le 2^{r_2-r_1} - 1, \\ &\textit{and} \quad 0 \le u \le 2^{n-r_2} - 1\} \end{aligned} \tag{4.117}$$

and

$$\begin{aligned} \mathcal{D}^2(L) = \{&\mathbf{E}_{i,j}, \mathbf{E}_{i,j+2^{n-r_1}} : i = u + t_1 2^{n-r_2}, j = u + t_2 2^{n-r_2}, \\ &0 \le t_1 < t_2 \le 2^{r_2-r_1} - 1 \quad \textit{and} \quad 0 \le u \le 2^{n-r_2} - 1\}. \end{aligned} \tag{4.118}$$

Consider the set $\overline{\mathcal{D}}(L)$ formed from the sets in equations (4.116), (4.117), and (4.118) by

$$\overline{\mathcal{D}}(L) = \mathbb{D}_2(L) \setminus (\mathcal{D}^1(L) \cup \mathcal{D}^2(L)). \tag{4.119}$$

Then the sets $\mathcal{A}(L)$, $\mathcal{A}(L) + \mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{D}_1(L)$, and $\mathcal{A}(L) + \mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{D}}(L)$, are pairwise disjoint and constitute all of $\mathcal{A}_2(L)$. Furthermore,

$$\mathcal{N}_2(L) = \left(\binom{2^{n-r_1+1}}{2} - 2^{n-r_2}(2^{2r_2-2r_1} - 1) + 2^{n-r_1+1} + 1 \right) 2^{L-1}. \tag{4.120}$$

Proof. Note that any L such that $w_H(2^n - L) \ge 3$ can be expressed as in equation (4.114). The characterization in equation (4.115) follows by using $r = 2$ in the hypothesis of Theorem 4.29 and $k = 2$ in equation (4.108). The rest of the proof deals with determining the disjoint set decomposition of $\mathcal{A}_2(L)$ in equation (4.115) from which we obtain the expression for $\mathcal{N}_2(L)$.

Case 1: $r_1 = r_2 = 1$
When $r_1 = r_2 = 1$ we have $1 \le L < 2^{n-2}$ and the characterization and counting function are already covered by Theorem 4.29(ii) with $r = 2$ and $k = 2$ in equation (4.109). Also, note that the expression for the counting function in equation (4.109) with $k = 2$ equals that in equation (4.120) with $r_1 = r_2 = 1$.

Case 2: $1 = r_1 < r_2$ or $1 < r_1 \le r_2$
First we determine the disjoint sets in $\mathcal{A}(L)[\mathbb{E}_1]$. By equation (4.114) we have

$$2^n - 2^{n-r_1+1} < L < 2^n - 2^{n-r_1}. \tag{4.121}$$

Using Theorem 4.27 and Lemma 4.8, from equation (4.121) we have

$$(\mathcal{A}(L)+\mathbf{E}_u)\cap(\mathcal{A}(L)+\mathbf{E}_v)=\emptyset,\quad 0\le u<v\le 2^{n-r_1+1}-1, \tag{4.122}$$

and for $u=0,\ldots,2^{n-r_1+1}-1$,

$$\mathcal{A}(L)+\mathbf{E}_u=\mathcal{A}(L)+\mathbf{E}_{u+t2^{n-r_1+1}},\quad t=0,\ldots,2^{r_1-1}-1. \tag{4.123}$$

Thus, from equation (4.122) there are 2^{n-r_1+1} disjoint sets $\mathcal{A}(L)+\mathbf{E}_i$, $\mathbf{E}_i\in\mathbb{D}_1(L)$, in $\mathcal{A}(L)[\mathbb{E}_1]$. To obtain the disjoint sets in $\mathcal{A}(L)[\mathbb{E}_2]$, we only have to characterize the disjoint sets in $\mathcal{A}(L)[\mathbb{D}_2(L)]$ because from equation (4.123) we have $\mathcal{A}(L)+\mathbf{E}_{i,j,i+v2^{n-r_1+1},j+w2^{n-r_1+1}}=\mathcal{A}(L)$, for $0\le i<j\le 2^{n-r_1+1}-1$ and $0\le v,w\le 2^{r_1-1}-1$.

From Theorem 4.27, we know that $\mathcal{A}(L)+\mathbf{E}_{i,j}=\mathcal{A}(L)+\mathbf{E}_{k,l}$ if and only if there exists a sequence $\mathbf{S}\in\mathcal{A}(L)$ such that $\mathbf{S}+\mathbf{E}_{i,j,k,l}\in\mathcal{A}(L)$. Hence we observe that redundantly counted sets in $\mathcal{A}(L)[\mathbb{D}_2(L)]$ arise if and only if there exist integers i, j, k, and l, $0\le i<j<k<l\le 2^{n-r_1+1}-1$, that are in the form given in equation (4.66). So the sets of integers i, j, k, and l, $0\le i<j<k<l\le 2^{n-r_1+1}-1$, such that $L(\mathbf{S}_{i,j,k,l})=L(\mathbf{S})$ for any $\mathbf{S}\in\mathcal{A}(L)$ are thus the i, j, k, and l in the form

$$i=u+g_1 2^{n-r_2},\quad j=u+g_2 2^{n-r_2},\quad k=i+2^{n-r_1},\quad l=j+2^{n-r_1}, \tag{4.124}$$

where

$$0\le u\le 2^{n-r_2}-1\quad\text{and}\quad 0\le g_1<g_2\le 2^{r_2-r_1}-1. \tag{4.125}$$

So for all settings of i and j in equation (4.124) we have the set equalities

$$\mathcal{A}(L)+\mathbf{E}_{i,j}=\mathcal{A}(L)+\mathbf{E}_{i+2^{n-r_1},j+2^{n-r_1}} \tag{4.126}$$

and

$$\mathcal{A}(L)+\mathbf{E}_{i,j+2^{n-r_1}}=\mathcal{A}(L)+\mathbf{E}_{i+2^{n-r_1},j}. \tag{4.127}$$

Also, for each $u=0,\ldots,2^{n-r_2}-1$, we have $2^{r_2-r_1}-1$ set equalities

$$\mathcal{A}(L)+\mathbf{E}_{u,u+2^{n-r_1}}=\mathcal{A}(L)+\mathbf{E}_{i,i+2^{n-r_1}},\quad\text{where}\quad i=u+t2^{n-r_2} \tag{4.128}$$

for $1\le t\le 2^{r_2-r_1}-1$.

Note that each error vector appearing on the left hand side or right hand side of equations (4.126) or (4.127) corresponding to all settings of i and j in equation (4.124) appears in only one of those equations and does not appear in the set equalities in equation (4.128). Also note that each error vector appearing on the left hand side or right hand side of set equalities in equation (4.128) does not appear in left hand side or right hand side of equations (4.126) and (4.127). Thus by equation (4.124), each of the set equalities in equations (4.126) and (4.127) results in a redundantly counted set in $\mathcal{A}(L)[\mathbb{D}_2(L)]$. These redundantly counted sets for all settings of i and j in equation (4.124) are listed as $\mathcal{A}(L)+\mathbf{E}_{i,j}$, $\mathbf{E}_{i,j}\in\mathcal{D}^2(L)$. Similarly, for each $u=0,\ldots,2^{n-r_2}-1$, the set equalities in equation (4.128) result in $2^{r_2-r_1}-1$ redundantly counted sets in $\mathcal{A}(L)[\mathbb{D}_2(L)]$. These redundantly counted sets are listed as $\mathcal{A}(L)+\mathbf{E}_{i,j}$, $\mathbf{E}_{i,j}\in\mathcal{D}^1(L)$.

Note that any L such that $2^{n-1} \leq L < 2^n$ and $w_H(2^n - L) \geq 3$, satisfies equations (4.111) and (4.113). From Lemma 4.8 and equation (4.121) we have

$$\mathcal{A}(L) \cap (\mathcal{A}(L) + \mathbf{E}_{i,j}) = \emptyset, \quad \mathbf{E}_{i,j} \in \mathbb{D}_2(L). \tag{4.129}$$

Thus, from equations (4.115), (4.111)–(4.113), (4.119), (4.124)–(4.128), and (4.129), the sets $\mathcal{A}(L)$, $\mathcal{A}(L) + \mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{D}_1(L)$, and $\mathcal{A}(L) + \mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{D}}(L)$, are mutually disjoint and constitute all of $\mathcal{A}_2(L)$.

From equations (4.117) and (4.118) we get

$$|\mathcal{D}^1(L)| = 2^{n-r_2}(2^{r_2-r_1} - 1) \quad \text{and} \quad |\mathcal{D}^2(L)| = 2^{n-r_2}\left(2\binom{2^{r_2-r_1}}{2}\right). \tag{4.130}$$

The number of disjoint sets in $\mathcal{A}(L)[\mathbb{E}_2]$ is equal to $|\overline{\mathcal{D}}(L)|$. From equations (4.119) and (4.130) we have

$$\begin{aligned} |\overline{\mathcal{D}}(L)| &= |\mathbb{D}_2(L)| - (|\mathcal{D}^1(L)| + |\mathcal{D}^2(L)|) \\ &= \binom{2^{n-r_1+1}}{2} - 2^{n-r_2}\left(2^{r_2-r_1} - 1 + 2\binom{2^{r_2-r_1}}{2}\right). \end{aligned} \tag{4.131}$$

From Lemma 4.2 we have $|\mathcal{A}(L)| = 2^{L-1}$, $1 \leq L \leq 2^n$. Hence the counting function in equation (4.120) follows from equations (4.115), (4.111)–(4.113), (4.122), (4.129), and (4.131). This completes the proof of the theorem. □

Next we give the characterization of 2^n-periodic binary sequences with fixed 3-error linear complexity L when $w_H(2^n - L) \neq 2$. Using the characterization we also obtain the corresponding counting function. For convenience we use the notation established in the statement of Theorem 4.30.

It is straightforward to see that

$$\mathcal{A}_3(0) = \mathbb{E}_1 \cup \mathbb{E}_2 \cup \mathbb{E}_3 \cup \{\mathbf{0}\} \quad \text{and} \quad \mathcal{N}_3(0) = \binom{2^n}{3} + \binom{2^n}{2} + 2^n + 1.$$

We also have $\mathcal{A}_3(2^n) = \emptyset$ and $\mathcal{N}_3(2^n) = 0$. From Lemma 4.26 we also get $\mathcal{A}_3(L) = \emptyset$ and $\mathcal{N}_3(L) = 0$ for $L = 2^n - 2^t$, $0 \leq t < n$.

Theorem 4.31. *Let $1 \leq L < 2^n$ be a positive integer such that $w_H(2^n - L) \geq 3$. Then*

$$\mathcal{A}_3(L) = \mathcal{A}_2(L) \bigcup \left(\bigcup_{\mathbf{E}_{i,j,k} \in \mathbb{E}_3} (\mathcal{A}(L) + \mathbf{E}_{i,j,k}) \right). \tag{4.132}$$

Furthermore, let L be bounded as

$$2^n - (2^{n-r_1} + 2^{n-r_2}) < L < 2^n - (2^{n-r_1} + 2^{n-r_2-1}),$$

for some r_1 and r_2 satisfying $1 \leq r_1 \leq r_2 \leq n-1$. Let $\overline{\mathcal{D}}(L)$ be as in equation (4.119). Define the sets $\mathbb{D}_3(L)$, $\mathcal{D}^3(L)$, and $\mathcal{E}(L)$ by

$$\mathbb{D}_3(L) = \{\mathbf{E}_{i,j,k} : 0 \leq i < j < k \leq 2^{n-r_1+1} - 1\},$$

$$\begin{aligned}\mathcal{D}^3(L) = \{&\mathbf{E}_{i,j,k}, \mathbf{E}_{i,j,l}, \mathbf{E}_{j,k,l}, \mathbf{E}_{i,k,l} : i = u + g_1 2^{n-r_2}, \quad j = u + g_2 2^{n-r_2}, \\ &k = i + 2^{n-r_1}, \quad l = j + 2^{n-r_1}, \quad 0 \le g_1 < g_2 < 2^{r_2-r_1}, \\ &\textit{and} \quad 0 \le u \le 2^{n-r_2} - 1\},\end{aligned} \tag{4.133}$$

and

$$\mathcal{E}(L) = \mathbb{D}_3(L) \setminus \mathcal{D}^3(L). \tag{4.134}$$

Then the sets $\mathcal{A}(L)$, $\mathcal{A}(L) + \mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{D}_1(L)$, $\mathcal{A}(L) + \mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{D}}(L)$, *and* $\mathcal{A}(L) + \mathbf{E}_{i,j,k}$, $\mathbf{E}_{i,j,k} \in \mathcal{E}(L)$, *are pairwise disjoint and constitute all of* $\mathcal{A}_3(L)$. *Furthermore,*

$$\mathcal{N}_3(L) = \mathcal{N}_2(L) + \left(\binom{2^{n-r_1+1}}{3} - 4 \cdot 2^{n-r_2}\binom{2^{r_2-r_1}}{2}\right) 2^{L-1}. \tag{4.135}$$

Proof. The characterization in equation (4.132) follows by using $r = 2$ in the hypothesis of Theorem 4.29 and $k = 3$ in equation (4.108). The rest of the proof deals with determining the disjoint set decomposition of $\mathcal{A}_3(L)$ in equation (4.132) from which we obtain the expression for $\mathcal{N}_3(L)$.

The case when $r_1 = r_2 = 1$, that is, when $1 \le L < 2^{n-2}$, is covered by Theorem 4.29(ii) with $r = 2$ and $k = 3$ in equation (4.109). It is straightforward to verify that the results using Theorem 4.29 when $r_1 = r_2 = 1$ agree with those stated in this theorem.

The rest of the proof handles the case when $r_1 = 1 < r_2$ or $1 < r_1 \le r_2$. We characterize the disjoint sets in the union given in equation (4.132). From Theorem 4.30 the disjoint sets in $\mathcal{A}_2(L)$ in equation (4.115) are $\mathcal{A}(L)$, $\mathcal{A}(L) + \mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{D}_1(L)$, and $\mathcal{A}(L) + \mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{D}}(L)$. Next we characterize the disjoint sets in $\mathcal{A}(L)[\mathbb{E}_3]$. For this, from equations (4.122) and (4.123) we only have to describe the disjoint sets in $\mathcal{A}(L)[\mathbb{D}_3(L)]$. From Theorem 4.21(ii) we can see that all sets in $\mathcal{A}(L)[\mathbb{D}_3(L)]$ are disjoint.

Finally, we show that the sets in $\mathcal{A}(L)[\mathbb{D}_3(L)]$ are disjoint from the sets $\mathcal{A}(L)$, $\mathcal{A}(L) + \mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{D}_1(L)$, and $\mathcal{A}(L) + \mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{D}}(L)$. Since the Hamming weight of any sequence in the sets in $\mathcal{A}(L)[\mathbb{D}_3(L)]$ is odd, these sets are disjoint from sets $\mathcal{A}(L)$ and $\mathcal{A}(L)+\mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{D}}(L)$. From Theorem 4.27, a set $\mathcal{A}(L)+\mathbf{E}_i$, $0 \le i \le 2^{n-r_1+1}-1$, is equal to some set $\mathcal{A}(L) + \mathbf{E}_{j,k,l}$, $0 \le j, k, l \le 2^{n-r_1+1} - 1$, if and only if there exists a sequence $\mathbf{S} \in \mathcal{A}(L)$ such that $\mathbf{S}_{i,j,k,l} \in \mathcal{A}(L)$. All such i, j, k, and l are described in equations (4.124) and (4.125). From equations (4.124) and (4.125), for each $u = 0, \ldots, 2^{n-r_2}-1$ there are exactly $\binom{2^{r_2-r_1}}{2}$ distinct pairs i, j and hence distinct sets $\{i, j, k, l\}$ such that $0 \le i < j < k < l \le 2^{n-r_1+1} - 1$ and $\mathcal{A}(L) + \mathbf{E}_{i,j,k,l} = \mathcal{A}(L)$. For each such distinct set $\{i, j, k, l\}$ we have four set equalities

$$\begin{aligned}&\mathcal{A}(L) + \mathbf{E}_{i,j,k} = \mathcal{A}(L) + \mathbf{E}_l, \quad \mathcal{A}(L) + \mathbf{E}_{i,j,l} = \mathcal{A}(L) + \mathbf{E}_k, \\ &\mathcal{A}(L) + \mathbf{E}_{j,k,l} = \mathcal{A}(L) + \mathbf{E}_i, \quad \text{and} \quad \mathcal{A}(L) + \mathbf{E}_{i,k,l} = \mathcal{A}(L) + \mathbf{E}_j.\end{aligned} \tag{4.136}$$

Based on the settings of possible i, j, k, and l in the equations we note that each error vector with Hamming weight 3 that appears in the set equalities in equation (4.136) appears in exactly one of them. This leads to four redundantly counted sets for each distinct setting of i, j, k, and l as described above. Thus all the redundantly

counted sets in the intersection of $\mathcal{A}(L)[\mathbb{D}_3(L)]$ and $\mathcal{A}(L)[\mathbb{D}_1(L)]$ are $\mathcal{A}(L)+\mathbf{E}_{i,j,k}$, $\mathbf{E}_{i,j,k} \in \mathcal{D}^3(L)$. Hence the sets in $\mathcal{E}(L)$ in equation (4.134) are disjoint from the sets $\mathcal{A}(L)$, $\mathcal{A}(L)+\mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{D}_1(L)$, and $\mathcal{A}(L)+\mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{D}}(L)$. Using the definition of k-error linear complexity the sets $\mathcal{A}(L)$, $\mathcal{A}(L)+\mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{D}_1(L)$, $\mathcal{A}(L)+\mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{D}}(L)$, and $\mathcal{A}(L)+\mathbf{E}_{i,j,k}$, $\mathbf{E}_{i,j,k} \in \mathcal{E}(L)$, are mutually disjoint and thus constitute all of $\mathcal{A}_3(L)$. Using this, the counting function in equation (4.135) follows from the definition of $\mathcal{E}(L)$ in equation (4.134). □

4.3.4 Characterization When $w_H(2^n - L) = 2$

We use results in Section 4.3.1 and the notation established in Section 4.3.3 to obtain the characterization of sequences in $\mathcal{A}(L)$ with fixed 2-error or 3-error linear complexity when $L = 2^n - (2^{n-r_1} + 2^{n-r_2})$, $1 \le r_1 < r_2 \le n$.

Theorem 4.32. *Let $L = 2^n - (2^{n-r_1} + 2^{n-r_2})$ for some $1 \le r_1 < r_2 \le n$. Define the sets*

$$\begin{aligned} \mathbb{G}_1(L) &= \{\mathbf{E}_i : 0 \le i \le 2^{n-r_1+1} - 1\} \quad \textit{and} \\ \mathbb{G}_2(L) &= \{\mathbf{E}_{i,j} : 0 \le i < j \le 2^{n-r_1+1} - 1\}. \end{aligned} \tag{4.137}$$

Consider the sets

$$\mathcal{H}^1(L) = \{\mathbf{E}_{i,i+2^{n-r_1}} : 0 \le i \le 2^{n-r_1} - 1\}, \tag{4.138}$$

$$\begin{aligned} \mathcal{H}^2(L) = \{&\mathbf{E}_{i,j}, \mathbf{E}_{i+2^{n-r_1},j+2^{n-r_1}}, \mathbf{E}_{i,j+2^{n-r_1}}, \mathbf{E}_{j,i+2^{n-r_1}} : \\ &0 \le i \le 2^{n-r_1} - 2^{n-r_2} - 1, \quad j = i + (2t+1)2^{n-r_2}, \\ &\textit{and} \quad 0 \le t \le 2^{r_2-r_1-1} - 1 - \lceil (\lfloor i/2^{n-r_2} \rfloor)/2 \rceil\}, \end{aligned} \tag{4.139}$$

and

$$\begin{aligned} \mathcal{H}^3(L) = \{&\mathbf{E}_{i,j}, \mathbf{E}_{i,j+2^{n-r_1}} : i \equiv j \bmod 2^{n-r_2+1} \\ &\textit{where} \quad 0 \le i < j < 2^{n-r_1}\}. \end{aligned} \tag{4.140}$$

Finally, define the set

$$\overline{\mathcal{H}}(L) = \mathbb{G}_2(L) \setminus (\mathcal{H}^1(L) \cup \mathcal{H}^2(L) \cup \mathcal{H}^3(L)). \tag{4.141}$$

Then the sets $\mathcal{A}(L)$, $\mathcal{A}(L) + \mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{G}_1(L)$, and $\mathcal{A}(L) + \mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{H}}(L)$, are pairwise disjoint and constitute all of $\mathcal{A}_2(L)$. That is

$$\mathcal{A}_2(L) = \mathcal{A}(L) \bigcup \left(\bigcup_{\mathbf{E}_i \in \mathbb{G}_1(L)} (\mathcal{A}(L) + \mathbf{E}_i) \right) \bigcup \left(\bigcup_{\mathbf{E}_{i,j} \in \overline{\mathcal{H}}(L)} (\mathcal{A}(L) + \mathbf{E}_{i,j}) \right). \tag{4.142}$$

Furthermore,

$$\mathcal{N}_2(L) = \left(\binom{2^{n-r_1+1}}{2} - 3 \cdot 2^{n+r_2-2r_1-1} + 2^{n-r_1+1} + 1 \right) 2^{L-1}. \tag{4.143}$$

Proof. By the definition of k-error linear complexity we have

$$\mathcal{A}_2(L) \subseteq \mathcal{A}(L) \bigcup \left(\bigcup_{\mathbf{E}_i \in \mathbb{E}_1} (\mathcal{A}(L) + \mathbf{E}_i) \right) \bigcup \left(\bigcup_{\mathbf{E}_{i,j} \in \mathbb{E}_2} (\mathcal{A}(L) + \mathbf{E}_{i,j}) \right). \tag{4.144}$$

For the rest of the proof let $\mathbf{S}$ be any sequence in $\mathcal{A}(L)$. By Lemma 4.1 we have $L_2(\mathbf{S}) = L$ and by Lemma 4.28 we get $L_2(\mathbf{S} + \mathbf{E}_i) = L$ for any $\mathbf{E}_i \in \mathbb{E}_1$. Thus

$$\mathcal{A}(L) \bigcup \left(\bigcup_{\mathbf{E}_i \in \mathbb{E}_1} (\mathcal{A}(L) + \mathbf{E}_i) \right) \subseteq \mathcal{A}_2(L). \tag{4.145}$$

Since $2^n - 2^{n-r_1+1} < L < 2^n - 2^{n-r_1}$, equations (4.122) and (4.123) also hold in the current setting. Thus there are 2^{n-r_1+1} disjoint sets $\mathcal{A}(L) + \mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{G}_1(L)$, in $\mathcal{A}(L)[\mathbb{E}_1]$. So we have

$$\bigcup_{\mathbf{E}_i \in \mathbb{E}_1} (\mathcal{A}(L) + \mathbf{E}_i) = \bigcup_{\mathbf{E}_i \in \mathbb{G}_1(L)} (\mathcal{A}(L) + \mathbf{E}_i). \tag{4.146}$$

Equations (4.122) and (4.123) also imply that $\mathcal{A}(L)[\mathbb{E}_2] = \mathcal{A}(L)[\mathbb{G}_2(L)]$. Next we determine which of the sets in $\mathcal{A}(L)[\mathbb{G}_2(L)]$ have sequences that belong to $\mathcal{A}_2(L)$. Equations (4.88) and (4.89) describe all distinct four symbol changes i, j, k, and l, $0 \le i < j < k < l \le 2^{n-r_1+1} - 1$, such that $L(\mathbf{S}_{i,j,k,l}) < L$. By equations (4.88) and (4.89) it is evident that for each integer u, $0 \le u \le 2^{n-r_1} - 1$, there exist integers v_1 and v_2, $0 \le v_1, v_2 \le 2^{n-r_1+1} - 1$, such that $L(\mathbf{S} + \mathbf{E}_{u,u+2^{n-r_1}} + \mathbf{E}_{v_1,v_2}) < L$. Thus

$$\forall\ \mathbf{S} \in \mathcal{A}(L) \quad \exists\ i,j: \ \text{If} \quad \mathbf{E}_{i,j} \in \mathcal{H}^1(L) \quad \text{then} \quad L_2(\mathbf{S} + \mathbf{E}_{i,j}) < L. \tag{4.147}$$

For each set of four symbol changes in equation (4.88) there are four distinct sequences $\mathbf{E}_{i,j}$, $\mathbf{E}_{i,j+2^{n-r_1}}$, $\mathbf{E}_{j,i+2^{n-r_1}}$, and $\mathbf{E}_{i+2^{n-r_1},j+2^{n-r_1}}$ in $\mathbb{G}_2(L)$ that when added to $\mathbf{S}$ result in sequences with 2-error linear complexity less than L. That is

$$\forall\ \mathbf{S} \in \mathcal{A}(L) \quad \exists\ i,j: \ \text{If} \quad \mathbf{E}_{i,j} \in \mathcal{H}^2(L) \quad \text{then} \quad L_2(\mathbf{S} + \mathbf{E}_{i,j}) < L. \tag{4.148}$$

By equations (4.147), (4.148), and Theorem 4.25(ii) we have

$$\forall\ \mathbf{S} \in \mathcal{A}(L) \quad \exists\ i,j: \ \text{If} \quad \mathbb{G}_2(L) \setminus (\mathcal{H}^1(L) \cup \mathcal{H}^2(L)) \quad \text{then} \quad L_2(\mathbf{S} + \mathbf{E}_{i,j}) = L$$

and thus

$$\begin{aligned} \bigcup_{\mathbf{E}_{i,j} \in \mathbb{G}_2(L) \setminus (\mathcal{H}^1(L) \cup \mathcal{H}^2(L))} (\mathcal{A}(L) + \mathbf{E}_{i,j}) \subseteq \mathcal{A}_2(L) \quad \text{and} \\ \bigcup_{\mathbf{E}_{i,j} \in \mathcal{H}^1(L) \cup \mathcal{H}^2(L)} (\mathcal{A}(L) + \mathbf{E}_{i,j}) \cap \mathcal{A}_2(L) = \emptyset. \end{aligned} \tag{4.149}$$

Next we describe the disjoint sets in $\{\mathcal{A}(L) + \mathbf{E}_{i,j} : \mathbf{E}_{i,j} \in \mathbb{G}_2(L) \setminus (\mathcal{H}^1(L) \cup \mathcal{H}^2(L))\}$. From Theorem 4.27, we know that $\mathcal{A}(L) + \mathbf{E}_{i,j} = \mathcal{A}(L) + \mathbf{E}_{k,l}$ if and only if there exists a sequence $\mathbf{R} \in \mathcal{A}(L)$ such that the new sequence $\mathbf{R} + \mathbf{E}_{i,j,k,l}$ is in $\mathcal{A}(L)$. Exactly all such i, j, k, and l are in the form given in equations (4.81) and (4.82). From the definitions in equations (4.138)–(4.140) we see the following.

(i) If $\mathbf{E}_{i,j} \in \mathcal{H}^1(L)$ then $j-i$ is 2^{n-r_1}.

(ii) If $\mathbf{E}_{i,j} \in \mathcal{H}^2(L)$ then $j-i$ is an odd multiple of 2^{n-r_2}.

(iii) If $\mathbf{E}_{i,j} \in \mathcal{H}^3(L)$ then $j-i$ is an even multiple of 2^{n-r_2} and $|j-i| < 2^{n-r_1}$.

From these observations we conclude

$$\mathcal{H}^{m_1}(L) \cap \mathcal{H}^{m_2}(L) = \emptyset, \quad 1 \le m_1 < m_2 \le 3. \tag{4.150}$$

For each of the $2^{n-r_2+1}\binom{2^{r_2-r_1-1}}{2}$ distinct settings of i and j in equations (4.81) and (4.82) the set equalities in equations (4.126) and (4.127) hold. By equation (4.150) and using an argument similar to that used in Theorem 4.30, this implies that there are $2 \cdot 2^{n-r_2+1}\binom{2^{r_2-r_1-1}}{2}$ redundantly counted sets in $\{\mathcal{A}(L) + \mathbf{E}_{i,j} : \mathbf{E}_{i,j} \in \mathbb{G}_2(L) \setminus (\mathcal{H}^1(L) \cup \mathcal{H}^2(L))\}$ enumerated as $\mathcal{A}(L) + \mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \mathcal{H}^3(L)$. (Note that the distinct settings of i and j in equations (4.81) and (4.82) are identical to those in equation (4.140).) So we have

$$\begin{aligned} &\bigcup_{\mathbf{E}_{i,j} \in \mathbb{G}_2(L) \setminus (\mathcal{H}^1(L) \cup \mathcal{H}^2(L))} (\mathcal{A}(L) + \mathbf{E}_{i,j}) \\ &= \bigcup_{\mathbf{E}_{i,j} \in \mathbb{G}_2(L) \setminus (\mathcal{H}^1(L) \cup \mathcal{H}^2(L) \cup \mathcal{H}^3(L))} (\mathcal{A}(L) + \mathbf{E}_{i,j}). \end{aligned} \tag{4.151}$$

Since $2^n - 2^{n-r_1+1} < L < 2^n - 2^{n-r_1}$, by Lemma 4.8 and Theorem 4.27 we have

$$\begin{aligned} \mathcal{A}(L) \cap (\mathcal{A}(L) + \mathbf{E}_u) &= \emptyset, \\ \mathcal{A}(L) \cap (\mathcal{A}(L) + \mathbf{E}_{i,j}) &= \emptyset, \quad \text{and} \\ (\mathcal{A}(L) + \mathbf{E}_u) \cap (\mathcal{A}(L) + \mathbf{E}_{i,j}) &= \emptyset, \end{aligned} \tag{4.152}$$

for all $\mathbf{E}_u \in \mathbb{G}_1(L)$ and $\mathbf{E}_{i,j} \in \mathbb{G}_2(L)$. Thus by equations (4.144)–(4.146) and (4.149)–(4.152) the sets $\mathcal{A}(L)$, $\mathcal{A}(L) + \mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{G}_1(L)$, and $\mathcal{A}(L) + \mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{H}}(L)$, are mutually disjoint and constitute all of $\mathcal{A}_2(L)$ and the characterization in equation (4.142) follows.

By equations (4.138) and (4.140) we have

$$\begin{aligned} &|\mathcal{H}^1(L)| = 2^{n-r_1} \quad \text{and} \\ &|\mathcal{H}^3(L)| = 2 \cdot 2^{n-r_2+1}\binom{2^{r_2-r_1-1}}{2} = 2^{n+r_2-2r_1-1} - 2^{n-r_1}. \end{aligned} \tag{4.153}$$

Each set of four symbol changes in equation (4.81) contributes four elements to the cardinality of $\mathcal{H}^2(L)$ as specified in equation (4.139). So by equations (4.90) and (4.139) we have

$$|\mathcal{H}^2(L)| = 4 \cdot 2^{n+r_2-2r_1-2} = 2^{n+r_2-2r_1}. \tag{4.154}$$

Thus by equations (4.137), (4.141), (4.150), (4.153), and (4.154) we obtain

$$\begin{aligned} |\overline{\mathcal{H}}(L)| &= |\mathbb{G}_2(L)| - (|\mathcal{H}^1(L)| + |\mathcal{H}^2(L)| + |\mathcal{H}^3(L)|) \\ &= \binom{2^{n-r_1+1}}{2} - \left(2^{n-r_1} + 2^{n+r_2-2r_1} + 2^{n+r_2-2r_1-1} - 2^{n-r_1}\right) \\ &= \binom{2^{n-r_1+1}}{2} - 3 \cdot 2^{n+r_2-2r_1-1}. \end{aligned} \tag{4.155}$$

The counting function in equation (4.143) follows from equations (4.3), (4.137), (4.142), and (4.155). □

For convenience, we use the notation established in the statement of Theorem 4.32 in the next result.

Theorem 4.33. *Let $L = 2^n - (2^{n-r_1} + 2^{n-r_2})$ for some $1 \leq r_1 < r_2 \leq n$. Define the sets $\mathbb{G}_3(L)$, $\mathcal{M}^1(L)$, and $\mathcal{M}^2(L)$ by*

$$\mathbb{G}_3(L) = \{\mathbf{E}_{i,j,k} : 0 \leq i < j < k \leq 2^{n-r_1+1}\},$$

$$\begin{aligned} \mathcal{M}^1(L) = \bigcup_{i=0}^{2^{n-r_1}-2^{n-r_2}-1} \{\mathbf{E}_{i,j,k}, \mathbf{E}_{i,j,l}, \mathbf{E}_{i,k,l}, \mathbf{E}_{j,k,l} : \quad & j = i + (2t+1)2^{n-r_2}, \\ & k = i + 2^{n-r_1}, \quad l = j + 2^{n-r_1}, \\ & \text{and} \quad 0 \leq t \leq 2^{r_2-r_1-1} - 1 - \lceil (\lfloor i/2^{n-r_2} \rfloor)/2 \rceil\}, \end{aligned} \tag{4.156}$$

and

$$\begin{aligned} \mathcal{M}^2(L) = \bigcup_{u=0}^{2^{n-r_2+1}-1} \{\mathbf{E}_{i,j,k}, \mathbf{E}_{i,j,l}, \mathbf{E}_{i,k,l}, \mathbf{E}_{j,k,l} : & \, i = u + g_1 2^{n-r_2+1}, \\ & j = u + g_2 2^{n-r_2+1}, \quad k = i + 2^{n-r_1}, \quad l = j + 2^{n-r_1}, \\ & \text{and} \quad 0 \leq g_1 < g_2 \leq 2^{r_2-r_1-1} - 1\}. \end{aligned} \tag{4.157}$$

Finally, define the set

$$\overline{\mathcal{M}}(L) = \mathbb{G}_3(L) \setminus (\mathcal{M}^1(L) \cup \mathcal{M}^2(L)). \tag{4.158}$$

Let $\overline{\mathcal{H}}(L)$ be as in equation (4.141) in Theorem 4.32. Then the sets $\mathcal{A}(L)$, $\mathcal{A}(L)+\mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{H}}(L)$, and $\mathcal{A}(L) + \mathbf{E}_{i,j,k}$, $\mathbf{E}_{i,j,k} \in \overline{\mathcal{M}}(L)$, are pairwise disjoint and constitute all of $\mathcal{A}_3(L)$. That is,

$$\begin{aligned} \mathcal{A}_3(L) = \mathcal{A}(L) \bigcup & \left(\bigcup_{\mathbf{E}_{i,j} \in \overline{\mathcal{H}}(L)} (\mathcal{A}(L) + \mathbf{E}_{i,j}) \right) \\ & \bigcup \left(\bigcup_{\mathbf{E}_{i,j,k} \in \overline{\mathcal{M}}(L)} (\mathcal{A}(L) + \mathbf{E}_{i,j,k}) \right). \end{aligned} \tag{4.159}$$

Furthermore,

$$\mathcal{N}_3(L) = \left(\binom{2^{n-r_1+1}}{3} + \binom{2^{n-r_1+1}}{2} - 7 \cdot 2^{n+r_2-2r_1-1} + 2^{n-r_1+1} + 1 \right) 2^{L-1}. \tag{4.160}$$

Proof. By the definition of k-error linear complexity we have

$$\mathcal{A}_3(L) \subseteq \mathcal{A}(L) \bigcup_{t=1}^{3} \left(\bigcup_{\mathbf{E}_{i_1,\ldots,i_t} \in \mathbb{E}_t} (\mathcal{A}(L) + \mathbf{E}_{i_1,\ldots,i_t}) \right). \tag{4.161}$$

For the rest of the proof let $\mathbf{S}$ be any sequence in $\mathcal{A}(L)$. By Lemma 4.1 we have $L_3(\mathbf{S}) = L$ and so

$$\mathcal{A}(L) \subseteq \mathcal{A}_3(L). \tag{4.162}$$

Since $2^n - 2^{n-r_1+1} < L < 2^n - 2^{n-r_1}$, equations (4.122) and (4.123) also hold in the current setting. Thus there are 2^{n-r_1+1} disjoint sets $\mathcal{A}(L) + \mathbf{E}_i$, $\mathbf{E}_i \in \mathbb{G}_1(L)$, in $\mathcal{A}(L)[\mathbb{E}_1]$ and thus equation (4.146) holds. By the format of four symbol changes that decrease the linear complexity of $\mathbf{S}$ given in equations (4.88) and (4.89), for each $i_1 = 0, \ldots, 2^{n-r_1+1} - 1$, there exist three integers i_2, i_3, and i_4 such that $L(\mathbf{S}_{i_1,i_2,i_3,i_4}) < L$, which implies

$$\bigcup_{\mathbf{E}_i \in \mathbb{G}_1(L)} (\mathcal{A}(L) + \mathbf{E}_i) \cap \mathcal{A}_3(L) = \emptyset. \tag{4.163}$$

By the proof of Theorem 4.32 we know that sequences in sets $\mathcal{A}(L)+\mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \mathbb{E}_2$, with 3-error linear complexity L are given by the disjoint union

$$\bigcup_{\mathbf{E}_{i,j} \in \overline{\mathcal{H}}(L)} (\mathcal{A}(L) + \mathbf{E}_{i,j}) \subseteq \mathcal{A}_3(L). \tag{4.164}$$

Equations (4.122) and (4.123) imply $\mathcal{A}(L)[\mathbb{E}_3] = \mathcal{A}(L)[\mathbb{G}_3(L)]$. So it is sufficient to determine the sequences in sets $\mathcal{A}(L)+\mathbf{E}_{i,j,k}$, $\mathbf{E}_{i,j,k} \in \mathbb{G}_3(L)$, that belong to $\mathcal{A}_3(L)$. For each set of four symbol changes in equation (4.88) there are four distinct sequences $\mathbf{E}_{i,j,k}$, $\mathbf{E}_{i,j,l}$, $\mathbf{E}_{i,k,l}$, and $\mathbf{E}_{j,k,l}$ in $\mathbb{G}_3(L)$ that when added to $\mathbf{S}$ result in sequences with 3-error linear complexity less than L. That is

$$\bigcup_{\mathbf{E}_{i,j,k} \in \mathcal{M}^1(L)} (\mathcal{A}(L) + \mathbf{E}_{i,j,k}) \cap \mathcal{A}_3(L) = \emptyset. \tag{4.165}$$

Equations (4.81) and (4.82) describe all i, j, k, and l, $0 \le i < j < k < l \le 2^{n-r_1+1}-1$, such that $L(\mathbf{S}_{i,j,k,l}) = L$. For each set of these four symbol changes we have four set equalities $\mathcal{A}(L) + \mathbf{E}_i = \mathcal{A}(L) + \mathbf{E}_{j,k,l}$, $\mathcal{A}(L) + \mathbf{E}_j = \mathcal{A}(L) + \mathbf{E}_{i,k,l}$, $\mathcal{A}(L) + \mathbf{E}_k = \mathcal{A}(L) + \mathbf{E}_{i,j,l}$, and $\mathcal{A}(L) + \mathbf{E}_l = \mathcal{A}(L) + \mathbf{E}_{i,j,k}$. By equation (4.163) this implies that

$$\bigcup_{\mathbf{E}_{i,j,k} \in \mathcal{M}^2(L)} (\mathcal{A}(L) + \mathbf{E}_{i,j,k}) \cap \mathcal{A}_3(L) = \emptyset. \tag{4.166}$$

By equation (4.156) for each $\mathbf{E}_{i,j,k} \in \mathcal{M}^1(L)$ we have either $i - j$, $j - k$, or $k - i$ is an odd multiple of 2^{n-r_2}. By equation (4.157) for each $\mathbf{E}_{i,j,k} \in \mathcal{M}^2(L)$ we have $i - j$, $j - k$, and $k - i$ are all even multiples of 2^{n-r_2}. From this we see that

$$\mathcal{M}^1(L) \cap \mathcal{M}^2(L) = \emptyset. \tag{4.167}$$

By equations (4.165), (4.166), and (4.167), Theorem 4.21(ii) and Theorem 4.25(iii), and using the fact that an odd number of changes to $\mathbf{S}$ results in a sequence with linear complexity 2^n, the sequences in the sets $\mathcal{A}(L) + \mathbf{E}_{i,j,k}$, $\mathbf{E}_{i,j,k} \in \mathbb{G}_3(L)$ with 3-error linear complexity L are given by the disjoint union

$$\bigcup_{\mathbf{E}_{i,j,k} \in \mathbb{G}_3(L) \setminus (\mathcal{M}^1(L) \cup \mathcal{M}^2(L))} (\mathcal{A}(L) + \mathbf{E}_{i,j,k}) \subseteq \mathcal{A}_3(L). \tag{4.168}$$

By equations (4.161)–(4.164), (4.168), and using the fact that odd numbers of changes to $\mathbf{S}$ result in sequences with linear complexity 2^n, the sets $\mathcal{A}(L)$, $\mathcal{A}(L)+\mathbf{E}_{i,j}$, $\mathbf{E}_{i,j} \in \overline{\mathcal{H}}(L)$, and $\mathcal{A}(L)+\mathbf{E}_{i,j,k}$, $\mathbf{E}_{i,j,k} \in \overline{\mathcal{M}}(L)$, are mutually disjoint and constitute all of $\mathcal{A}_3(L)$ and the characterization in equation (4.159) follows.

From equations (4.81), (4.88), (4.90), and (4.156)–(4.158) we have

$$\begin{aligned} |\overline{\mathcal{M}}(L)| &= |\mathbb{G}_3(L)| - (|\mathcal{M}^1(L)| + |\mathcal{M}^2(L)|) \\ &= \binom{2^{n-r_1+1}}{3} - \left(4 \cdot 2^{n+r_2-2r_1-2} + 4 \cdot 2^{n-r_2+1}\binom{2^{r_2-r_1-1}}{2}\right) \\ &= \binom{2^{n-r_1+1}}{3} + 2^{n-r_1+1} - 4 \cdot 2^{n+r_2-2r_1-1}. \end{aligned} \tag{4.169}$$

The counting function in equation (4.160) follows from equations (4.3), (4.155), (4.159), and (4.169). □

4.3.5 Concluding Remarks

In this section we used algebraic and combinatorial methods to characterize and count the number of 2^n-periodic binary sequences with fixed 2-error or 3-error linear complexity. Here we make some observations based on the counting functions derived in the previous two sections.

Let $\mathcal{N}_{\geq}(L)$, $0 \leq L \leq 2^n$, be the number of 2^n-periodic binary sequences with linear complexity at least L. From Lemma 4.2 we have

$$\mathcal{N}_{\geq}(L) = \left(\frac{2^{2^n-L+1}-1}{2^{2^n-L+1}}\right) 2^{2^n}. \tag{4.170}$$

Define $f_k(L)$, $1 \leq k \leq 2^n$, by

$$f_k(L) = \frac{\mathcal{N}_k(L)}{\mathcal{N}_{\geq}(L)}. \tag{4.171}$$

That is, $f_k(L)$ describes the proportion of sequences with k-error linear complexity L among sequences with linear complexity at least L. For cryptographic purposes we would like to have $f_k(L)$ as high as possible for large L and at least for small k.

By equations (4.120), (4.143), (4.170), and (4.171) after simplification we obtain

$$f_2(L) = \frac{2^{2n-2r_1+1} + 2^{n-r_1} + 2^{n-r_2} + 1 - 2^{n+r_2-2r_1}}{2^{2^n-L+1}-1} \tag{4.172}$$

when $2^n - (2^{n-r_1} + 2^{n-r_2}) < L < 2^n - (2^{n-r_1} + 2^{n-r_2-1})$ with $1 \leq r_1 \leq r_2 \leq n-1$ and

$$f_2(L) = \frac{2^{2n-2r_1+1} + 2^{n-r_1} + 1 - 3 \cdot 2^{n+r_2-2r_1-1}}{2^{2^n-L+1}-1} \tag{4.173}$$

when $L = 2^n - (2^{n-r_1} + 2^{n-r_2})$, $1 \leq r_1 < r_2 \leq n$. Using these formulae we find $f_2(L)$ for $L = 2^n-3$, 2^n-5, 2^n-6, and 2^n-7. When $L = 2^n-7$, we have $w_H(2^n-L) = 3$ and we can uniquely bound L as $2^n - (2^{n-r_1} + 2^{n-r_2}) < L < 2^n - (2^{n-r_1} + 2^{n-r_2-1})$

with $r_1 = r_2 = n-2$. Using $L = 2^n - 7$ and $r_1 = r_2 = n-2$ in equation (4.172) we have $f_2(2^n - 7) = 37/255 \approx 1/7$. When $L = 2^n - 3$, we have $w_H(2^n - L) = 2$ and $L = 2^n - (2^{n-r_1} + 2^{n-r_2})$ with $r_1 = n-1$ and $r_2 = n$. So we have $f_2(2^n - 3) = 5/15 = 1/3$ by equation (4.173). Similarly we obtain $f_2(2^n - 5) = 13/63 \approx 1/5$ and $f_2(2^n - 6) = 25/127 \approx 1/5$. Using equations (4.135), (4.160), (4.170), and (4.171) we also obtain corresponding values for $f_3(L)$. Using Theorem 4.19 we determine the corresponding values for $f_1(L)$. All these values are summarized in Table 4.1. Since the number of sequences with high linear complexity is large for 2^n-periodic

Table 4.1: $f_1(L)$, $f_2(L)$, and $f_3(L)$ for large L

L	$f_1(L)$	$f_2(L)$	$f_3(L)$
$2^n - 3$	$1/3$	$1/3$	$1/15$
$2^n - 5$	$1/7$	$13/63 \approx 1/5$	$37/63 \approx 1/2$
$2^n - 6$	$9/127 \approx 1/14$	$25/127 \approx 1/5$	$65/127 \approx 1/2$
$2^n - 7$	$9/255 \approx 1/28$	$37/255 \approx 1/7$	$93/255 \approx 1/3$

binary sequences, we see that a considerable number of sequences have high linear complexity and high 2-error or 3-error linear complexity.

Using the counting functions derived in this section, statistical properties like expected value and variance can also be considered for the 2-error or 3-error linear complexity of 2^n-periodic binary sequences. The resulting expressions for the expected values are quite complicated and unlikely to yield a simple closed form. However, estimates may be possible. Extension to p^n-periodic sequences over $\mathbb{F}_p$ can also be considered. Similar results for periodic sequences with arbitrary period or periods of other forms are desirable.

Remark 4.5. It is important to note here that 2^n-periodic sequences may not be used as key streams for stream ciphers. This is true especially if the attacker knows that the key streams have period 2^n. In this case knowing a segment of length $t \geq L$, the linear complexity, would enable the attacker to recover the whole sequence using the fact that $(1-x)^t$ is a characteristic polynomial for the sequence.

5 Further Research

The sequence complexity measures discussed in this thesis are based on LFSRs and FCSRs. One can similarly define measures based on any class of generators. The measure would simply be the length of the shortest generator of that class that generates the given sequence. We note that even though all eventually periodic sequences can be generated by LFSRs, it is important to study other types of generators with synthesis algorithms because the length of the shortest LFSR that generates a sequence might be very high compared to the length of the shortest generator of a different class. Thus a cryptanalyst would have more tools now to analyze the cryptographic strength of a sequence compared to when only LFSRs were studied. In this chapter we discuss future research directions based on feedback shift register (FSR) based sequence complexity measures. Then we outline ideas for using research on sequence complexity measures for design and analysis of stream ciphers.

5.1 FSR Based Sequence Complexity Measures

As mentioned earlier, counting functions for N-adic complexity measures do not exist in the literature and are comparatively difficult to derive because certain nice properties of polynomials over finite fields do not apply to N-adic representations of integers. The expected value of k-error N-adic complexity is also not determined yet. The N-adic analogs for most of the results on k-error linear complexity also do not exist at the time of this writing.

Arithmetic k-Error N-adic Complexity

Besides the measures based on substitutions, insertions, and deletions, a new complexity measure is introduced in the following definition.

Definition 5.1. The arithmetic k-error N-adic complexity, $\lambda^a_{N,k}(\mathbf{S})$, of a periodic sequence $\mathbf{S}$ is the lowest N-adic complexity achieved by modifying a single period of $\mathbf{S}$ by N-adic addition or subtraction with carry of an error vector of length equal to the period and with Hamming weight at most k.

$\lambda^a_{N,k}(\mathbf{S})$ differs from $\lambda_{N,k}(\mathbf{S})$ in that the elements in the error vector are substituted for the corresponding elements in the $\mathbf{S}$ in the latter case while the error vector is N-adically added/subtracted in the former.

Let $\mathbf{S} = (s_0, \ldots, s_{T-1})^\infty$ and let $\mathbf{S}(N)$ denote $s_0 + s_1 N + \cdots + s_{T-1}N^{T-1}$. Then the N-adic complexity of $\mathbf{S}$ is $\log_N((N^T - 1)/\gcd(N^T - 1, \mathbf{S}(N)))$. Now $\lambda^a_{N,k}(\mathbf{S})$ can also be defined as

$$\lambda^a_{N,k}(\mathbf{S}) = \min\left\{\lambda_N\left(-\frac{(\mathbf{S}(N) \pm m)}{N^T - 1}\right) : m \geq 0 \ \textit{and} \ w_H(m) \leq k\right\},$$

where $\lambda_N(\mathbf{S})$ denotes the N-adic complexity.

Example 5.1. Let $N = 2, T = 8$, $\mathbf{S} = (10110000)^\infty$, and $k = 1$. So $\mathbf{S}(2) = 13$. The rational representation of the corresponding 2-adic number is $-13/255$. By using all possible error vectors the lowest 2-adic complexity can be obtained by subtracting 64 from $\mathbf{S}(2)$. So here $m = 64$ and the rational representation of the new sequence is $-(13 - 64)/255 = 51/255 = 1/5$. Thus we have $\lambda^a_{2,k=1}(\mathbf{S}) = \log_2 5$. Here $\lambda_2(\mathbf{S}) = \log_2 255$ and $\lambda_{2,k=1}(\mathbf{S}) = \log_2 17$.

We make the following two observations about $\lambda^a_{N,k}(\mathbf{S})$ from Example 5.1.

(i). The inclusion of both addition and subtraction of the error vector in the definition of $\lambda^a_{N,k}(\mathbf{S})$ is important. In this example no other m with $w_H(m) = 1$ exists that can be added to decrease the complexity to $\log_2 5$.

(ii). We note that after the modification the sequence may not be strictly periodic. Nevertheless, the complexity is reduced to the minimum possible with that k.

We have the following result on the relationship between $\lambda^a_{2,k}(\mathbf{S})$ and $\lambda_{2,k}(\mathbf{S})$.

Lemma 5.1. *Let* $\mathbf{S}$ *be a periodic binary sequence with period* T. *Then* $\lambda^a_{2,k}(\mathbf{S}) \leq \lambda_{2,k}(\mathbf{S})$ *for* $k = 1$. *This relationship or* $\lambda^a_{2,k}(\mathbf{S}) \geq \lambda_{2,k}(\mathbf{S})$ *does not hold for* $k > 1$.

Proof. The number of error vectors possible is T if $k = 1$. From Definition 5.1 $\lambda^a_{2,k}(\mathbf{S})$ is the minimum among the 2-adic complexities of $2T$ new sequences corresponding to $\mathbf{S}^T(2) \pm 2^i$, $i = 0, ..., T$. In the case of $\lambda_{2,k}(\mathbf{S})$, we note that XORing a period with an error vector with a 1 in the i^{th} position is equivalent to the 2-adic addition (resp. subtraction) if the corresponding bit in the period is a 0 (resp. 1) resulting in T different new sequences. Hence the set of these T new sequences is a subset of the set of $2T$ sequences considered to obtain $\lambda^a_{2,k}(\mathbf{S})$. The relationship does not hold with $k > 1$ because then the new sequences produced by XORing are not necessarily produced by 2-adic addition or subtraction. We show this using the following examples:

(i). $\lambda_{2,k}(\mathbf{S}) < \lambda^a_{2,k}(\mathbf{S})$: Let $T = 8$, $k = 2$, and $\mathbf{S} = (10001110)^\infty$, so that $S^T(2) = 113$. The rational representation of $\mathbf{S}$ is $-113/255$. The lowest $\lambda_2(\mathbf{S})$ can be obtained by adding 40 to $\mathbf{S}^T(2)$, and hence the rational representation of the new sequence is $-(113 + 40)/255 = -153/255 = -3/5$. So $\lambda^a_{2,k}(\mathbf{S}) = \log_2 5$. But for $\lambda_{2,k}(\mathbf{S})$ the lowest $\lambda_2(\mathbf{S})$ can be obtained by XORing with (00100100). We have $(10001110) \oplus (00100100) = (10101010) = 85$. So the rational representation of the new sequence is $-85/255 = -1/3$, and so $\lambda_{2,k}(\mathbf{S}) = \log_2 3$.

(ii). $\lambda^a_{2,k}(\mathbf{S}) < \lambda_{2,k}(\mathbf{S})$: Let $T = 8$, $k = 2$, and $\mathbf{S} = (01100110)^\infty$, so that $S^T(2) = 102$. The rational representation of $\mathbf{S}$ is $-102/255$. The lowest $\lambda_2(\mathbf{S})$ can be obtained by adding 68 to $\mathbf{S}^T(2)$, and hence the rational representation of the new sequence is $-(102 + 68)/255 = -170/255 = -2/3$. So $\lambda^a_{2,k}(\mathbf{S}) = \log_2 3$. But for $\lambda_{2,k}(\mathbf{S})$ the lowest $\lambda_2(\mathbf{S})$ can be obtained by XORing with (00000000). We have $(01100110) \oplus (00000000) = (01100110) = 102$. So the rational representation of the new sequence is $-102/255 = -2/5$, and so $\lambda_{2,k}(\mathbf{S}) = \log_2 5$.

□

A good starting point in analyzing the arithmetic k-error N-adic complexity would be to compute its expected value for periodic sequences and study its relationship with k-error N-adic complexity for $N > 2$.

Joint N-adic Complexity

Recall that the $\mathbb{F}_q$-linear complexity of an m-fold multisequence over $\mathbb{F}_q$ is the linear complexity of its corresponding single sequence over $\mathbb{F}_{q^m}$. We also know that the smallest LFSR with taps in $\mathbb{F}_q$ that generates the sequence over $\mathbb{F}_{q^m}$ also generates the m component sequences. Hence the $\mathbb{F}_q$-linear complexity is greater than or equal to the joint linear complexity. This follows from the fact that the corresponding recurrence over $\mathbb{F}_q$ still holds for all the m sequences over $\mathbb{F}_q$. Analogous to $\mathbb{F}_q$-linear complexity we consider the following complexity measure in the N-adic case.

Definition 5.2. The N-tap N^m-adic complexity, $\lambda_N^{N^m}(\mathbf{S})$, of a sequence $\mathbf{S}$ over $\{0, \ldots, N^m - 1\}$ is the size of the smallest FCSR with taps in $\{0, \ldots, N - 1\}$ that can generate $\mathbf{S}$.

We note that in Definition 5.2 if $\mathbf{S}$ is a periodic sequence, then an FCSR with taps in $\{0, \ldots, N-1\}$ always exists, for instance, with the initial memory and all tap coefficients except the one on the output cell set to 0.

Let $\mathbb{S}$ be an m-fold multisequence over $\{0, \ldots, N - 1\}$. We can identify the m component sequences of $\mathbb{S}$ as a single sequence $\mathcal{S}$ over $\{0, \ldots, N^m - 1\}$. But unlike in the linear complexity case, it is apparent that the smallest FCSR with taps in $\{0, \ldots, N - 1\}$ that generates the sequence over $\{0, \ldots, N^m - 1\}$ may not generate the component sequences. Hence there is no clear relationship between $\lambda_N(\mathbb{S})$ and $\lambda_N^{N^m}(\mathcal{S})$. Hong et al. [28] introduced a method of constructing sequences over $\mathbb{F}_{p^m}$ with characteristic polynomial over $\mathbb{F}_p$, which can be investigated in the N-adic case. A first task is to investigate the relationship between the joint N-adic complexity of m-fold multisequences over $\{0, \ldots, N - 1\}$ and the N-tap N^m-adic complexity of the corresponding single sequences over $\{0, \ldots, N^m - 1\}$. It might be helpful to find and evaluate new approaches to understand the joint N-adic complexity of multisequences. While there are multisequence LFSR synthesis algorithms [84, 87], similar algorithms are desirable for multisequence FCSR synthesis.

AFSRs and π-adic Complexity

AFSRs are generalizations of FCSRs and generate sequences over arbitrary finite fields [42]. The architecture of an AFSR is similar to that of an FCSR. Analogous to power series for LFSRs and N-adic numbers for FCSRs, π-adic numbers are used to analyze AFSRs. While the coefficients for power series and N-adic numbers are taken, respectively, from the underlying ring and the set $\{0, \ldots, N - 1\}$, there may be no such natural set to construct π-adic numbers. The expected value of the π-adic complexity of sequences was studied for only a few families of AFSRs [40]. A challenging problem is to extend these results to several other families of AFSRs.

NFSRs and t-th Order Complexity

A nonlinear feedback shift register (NFSR) is like an LFSR except that the feedback function is nonlinear as a function of the cells of the register. Associated with these registers we have *maximal order* complexity (or nonlinear span) of a sequence which is the length of the shortest NFSR that can generate the sequence. Considering efficiency as a key requirement, these NFSRs are generally useful only when the degree of the nonlinearity of the feedback function is small. This gives rise to the t-th order nonlinear complexity, which is the length of the shortest NFSR with the degree of the feedback function at most t that generates the sequence. Algorithms to find maximal order complexity of sequences have been studied recently [46, 73, 74]. Rizomiliotis gave some constructions of sequences with maximal nonlinear span [71] and provided cryptanalytic motivation for research on nonlinear complexity [72]. Counting functions, expected values, and efficient algorithms for the t-th order complexity for small t such as $t = 2$, 3 are desirable.

Sequences with Large Complexity

While there are some recent results for constructions of sequences with large linear complexity [29], similar constructions for the N-adic case are desirable. Binary sequences can be viewed as LFSR (resp. FCSR) sequences over $\mathbb{F}_2$ and also as sequences over any finite field (resp. $\{0, \ldots, N-1\}$). A useful task is to characterize binary sequences with high complexity over $\mathbb{F}_2$ but with significantly lower (higher) complexity when considered over other larger sets. Also, another interesting approach is to consider consecutive blocks of sequence bits of a fixed size m and treat each block as an element over $\mathbb{F}_{2^m}$ (resp. $\{0, \ldots, 2^m - 1\}$). If these new sequences have low linear (resp. 2^m-adic) complexity, then the original sequences become vulnerable. Here we would like to characterize binary sequences with high complexity that have the property that the sequences formed by consecutive blocks of sizes m have significantly lower (higher) complexities for small m. These results impact both the cryptanalysis and the design of cryptographically strong building blocks for stream ciphers.

Summary

Table 5.1 lists various problems dealing with FSR related sequence complexity measures. The right hand column lists various sequence complexity related problems and the left hand column indicates the different settings in which we can try to solve them. As surveyed in the earlier chapters, while several results are available in the linear case, results in the N-adic and other nonlinear cases are not known for many of these problems.

5.2 Design and Cryptanalysis of Stream Ciphers

In this section we briefly describe design principles and cryptanalytic techniques for stream ciphers. We propose two future research directions in the design and cryptanalysis of stream ciphers.

Table 5.1: Complexity measures and associated problems

Settings	Problems
FSR type: LFSR, FCSR, AFSR, or NFSR	Counting functions
Single sequences or Multisequences	Expected values
Periodic or Finite Length	Complexity bounds
Complexity type: Conventional, k-error, k-insert	Complexity computation
k-delete, k-operation, t-th order, arithmetic k-error	Shift register synthesis
Special Cases: $k = 1, 2$; $t = 2, 3$; $T = p^d$; Fix $\mathbb{F}_q$, N	Asymptotic behavior

Stream Cipher Standards

Owing to the high speed requirements and the ease in implementation, hardware implementations of stream ciphers based on LFSRs became popular in 1970s and 1980s. The A5 series of stream ciphers used in GSM cellular standard and the E0 stream cipher used in the short-range wireless radio standard Bluetooth are examples of such ciphers implemented in hardware. Word oriented software stream ciphers implemented in software were proposed in the 1990s; LEVIATHAN (Cisco), MUGI (Hitachi-K.U. Leuven), RC4 (R. Rivest), SNOW (Lund University), SOBER (Qualcomm), and SEAL (IBM) are a few examples of such ciphers. For encryption using small devices with limited resources, software based ciphers are slower and consume more energy compared to those implemented using dedicated hardware co-processors.

The eSTREAM Project

The NESSIE (New European Schemes for Signatures, Integrity, and Encryption) project ran from 2000 to 2004 to put together a portfolio of cryptographic primitives through an open and transparent process. Unfortunately no stream cipher made it to the final portfolio of NESSIE, as weaknesses were discovered in all proposed stream ciphers. In 2004 ECRYPT, a Network of Excellence within the Information Societies Technology (IST) Program of the European Commission was launched. In April 2005, ECRYPT received 34 candidate stream ciphers for its eSTREAM project. The final portfolio of eSTREAM consists of four software based stream ciphers and three implemented in hardware. Though the project concluded in 2008, research is still being conducted on the final candidates.

Types of Stream Ciphers

To ensure proper decryption the sender and the receiver must be synchronized. That is, they should be operating at the same position within the key. Synchronous stream ciphers use key streams produced independently of the plain text that is being encrypted. They need additional mechanisms to guarantee synchronization. In these

schemes, if a symbol is modified due to a transmission error, only that symbol will be decrypted erroneously. By definition synchronous ciphers guarantee no error propagation. Self-synchronizing stream ciphers suffer limited error propagation but can self-synchronize to recover from dropped or inserted symbols during the transmission. In these ciphers, the key stream is usually produced as a function of the key and a fixed number of previous cipher text symbols. Moustique and SSS are examples of eSTREAM candidates that are self-synchronizing. Only three out of the 34 submissions to eSTREAM belong to this category. It should be noted here that since these ciphers use cipher text to update the state, they should also withstand chosen plain text and chosen cipher text attacks while for synchronous ciphers one only has to worry about known plain text attacks. The eSTREAM final portfolio report declares the design of secure self-synchronizing stream ciphers a "very significant" open problem.

To deal with the inherent linearity of LFSRs several techniques were proposed [76]. In a *combination generator* output symbols from several LFSRs are combined using a nonlinear combining function, to obtain one key stream symbol at every clock. A different approach called the *filter generator* uses a nonlinear function to combine the symbols of the state of a single LFSR at every clock to produce a key stream symbol. Sfinks and WG are eSTREAM candidates that belong to this category.

Another recently pursued approach is to use nonlinear state updates. This can be done by clocking different LFSRs at different intervals to be used to determine key stream symbols. eSTREAM candidates Decim, Mickey, and Pomaranch are based on LFSRs with clocking. FCSRs, AFSRs, and NFSRs can also be used for nonlinear updates. FFCSR-H, DRAGON, NLS, and SSS are example submissions to eSTREAM that belong to this category.

Two Cryptanalytic Techniques

According to the eSTREAM final portfolio report some cryptanalysts currently feel that the security analysis of stream ciphers is somewhat *ad hoc*. There are a number of approaches that can be used for cryptanalysis including time-memory-data tradeoff attacks, guess and determine attacks, and side-channel attacks. Here we only mention two of the most important techniques.

A *correlation attack* [62, 79] exploits the correlation between the output of the nonlinear function used in a filter generator and the output of one of the LFSRs of the generator. Once high correlation is observed for a particular LFSR, using a brute force search on the initial states of the LFSR one can choose a state that produces that LFSR output with high correlation to the nonlinear function output. This task can be repeated for each individual LFSR involved. Note that the attacks are probabilistic in the sense that we *guess* the internal state that has a given correlation with the output key stream; this might not always give the correct internal state.

Recently, a new cryptanalytic technique called *algebraic cryptanalysis* [8, 9] was introduced. Algebraic attacks have cryptanalytic ramifications in all cryptographic primitives including block ciphers, public key primitives, hash functions, and stream ciphers. Algebraic attacks involve two steps:

(i). Find a set of "simple" multivariate equations that describe the cryptosystem.

(ii). Solve the system of equations obtained in step (i).

In the context of stream ciphers the multivariate equations involve the initial state (key) and the key stream. If the system of equations is sparse, low degree, and overdefined it will be easier to solve. Several techniques such as linearization, algorithms using Gröbner bases such as F4, and dedicated algorithms such as XL and XSL are available to solve such systems of equations. Although the time complexity of launching algebraic attacks is not completely understood, they are very interesting and useful because they require very few known plain texts.

Future Research Problems

In the context of stream cipher design and analysis we propose these two problems.

(i). Although most current stream cipher designs guarantee high complexities for key streams generated, complexities after a few modifications are performed on the key sequences are not studied thus far. Current literature on error complexity measures does not address their impact on the cryptanalysis of specific stream cipher designs. We plan to study the error complexity values for key streams used in stream ciphers selected for the eSTREAM portfolio to see if they expose any further weaknesses in them.

(ii). While there are several stream ciphers based on LFSRs, FFCSR stream ciphers are the first popular FCSR based ciphers featuring Galois mode FCSRs [22] and a linear filtering function. Although they were initially included in the eSTREAM final portfolio, a real time attack exploiting a weakness when all carry cells except the initial one contain zeroes was recently demonstrated [27]. Despite this specific attack, the memory component in FCSRs makes it difficult to launch algebraic and correlation attacks against them in general as the number and degree of multivariate equations would be very high. So an important research task is to design and analyze new FCSR based stream ciphers. As a first step we would like to explore whether Galois mode d-FCSRs [22] offer resistance to the real time attacks on FFCSR stream ciphers. Evaluating the potential of generalized AFSRs for use in stream ciphers is also a future challenging problem.

References

[1] A. Alecu and A. Sălăgean. A genetic algorithm for computing the k-error linear complexity of cryptographic sequences. In *IEEE Congress on Evolutionary Computation*, pages 3569–3576. IEEE, 2007. 37

[2] A. Alecu and A. Sălăgean. Modified Berlekamp-Massey algorithm for approximating the k-error linear complexity of binary sequences. In S. D. Galbraith, editor, *IMA Int. Conf.*, volume 4887 of *LNCS*, pages 220–232. Springer, 2007. 37

[3] F. Arnault and T. P. Berger. Design and properties of a new pseudorandom generator based on a filtered FCSR automaton. *IEEE Trans. on Computers*, 64(11):1364–1383, 2005. 2

[4] F. Arnault, T. P. Berger, and A. Necer. Feedback with carry shift registers synthesis with the euclidean algorithm. *IEEE Trans. Inform. Theory*, 50(5):910–917, 2004. 10

[5] S. Blackburn. Fast rational interpolation, reed-solomon decoding, and the linear complexity profiles of sequences. *IEEE Trans. Inform. Theory*, 43(2):537–548, 1997. 8

[6] R. Blahut. Transform techniques for error control codes. *IBM Journal of Research and Development*, 23:299–315, 1979. 16

[7] H. Chen. Fast algorithms for determining the linear complexity of sequences over $GF(p^m)$ with period $2^t n$. *IEEE Trans. Inform. Theory*, 51(5):1854–1856, 2005. 37

[8] N. Courtois. Fast algebraic attacks on stream ciphers with linear feedback. In D. Boneh, editor, *CRYPTO*, volume 2729 of *LNCS*, pages 176–194, 2003. 84

[9] N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. In E. Biham, editor, *EUROCRYPT*, volume 2656 of *LNCS*, pages 345–359, 2003. 84

[10] R. Couture and P. L'Ecuyer. Distribution properties of multiply-with-carry random number generators. *Mathematics of Computation*, 66:591–607, 1997. 5

[11] Z. Dai, S. Jiang, K. Imamura, and G. Gong. Asymptotic behavior of normalized linear complexity of ultimately nonperiodic binary sequences. *IEEE Trans. Inform. Theory*, 50(11):2911–2915, 2004. 11

[12] C. Ding, G. Xiao, and W. Shan. *The Stability Theory of Stream Ciphers*. Springer, 1991. 11, 37

[13] L. Dong, Y. Hu, and Y. Zeng. Computing the k-error N-adic complexity of a sequence of period p^n. In G. Gong, T. Helleseth, H.-Y. Song, and K. Yang, editors, *SETA 2006*, volume 4086 of *LNCS*, pages 190–198. Springer, 2006. 38

[14] Ecrypt stream ciphers project. http://www.ecrypt.eu.org/stream/. 12

[15] T. Etzion. Constructions for perfect maps and pseudorandom arrays. *IEEE Trans. Inform. Theory*, 34(5):1308–1316, 1988. 38

[16] T. Etzion, N. Kalouptsidis, N. Kolokotronis, K. Limniotis, and K. G. Patterson. On the error linear complexity profiles of binary sequences of period 2^n. *Int. Symp. Inform. Theory*, pages 2400–2404, 2008. 56, 57

[17] Z. Fengxiang and Q. Wenfeng. The 2-error linear complexity of 2^n-periodic binary sequences with linear complexity $2^n - 1$. *Journal of Electronics (China)*, 24(3):390–395, 2007. 45, 46, 52

[18] F.-W. Fu, H. Niederreiter, and M. Su. The characterization of 2^n-periodic binary sequences with fixed 1-error linear complexity. In G. Gong, T. Helleseth, H.-Y. Song, and K. Yang, editors, *SETA 2006*, volume 4086 of *LNCS*, pages 88–103. Springer, 2006. 44, 57, 58, 66

[19] R. A. Games and A. H. Chan. A fast algorithm for determining the complexity of a pseudo-random sequence with period 2^n. *IEEE Trans. Inform. Theory*, 29(1):144–146, 1983. 37, 40

[20] S. Golomb. *Shift Register Sequences*. Aegean Park Press, 1982. 4

[21] S. Golomb and G. Gong. *Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar*. Cambridge University Press, 2005. 4

[22] M. Goresky and A. Klapper. Fibonacci and Galois representations of feedback-with-carry shift registers. *IEEE Trans. Inform. Theory*, 48(11):2826–2836, 2002. 85

[23] M. Goresky and A. Klapper. *Algebraic Shift Register Sequences*. Preprint, 2007. 4, 7, 9, 10, 17

[24] M. Goresky, A. Klapper, and L. Washington. Fourier transforms and the 2-adic span of periodic binary sequences. *IEEE Trans. Inform. Theory*, 46(2):687–691, 2000. 16

[25] F. Gustavson. Analysis of the Berlekamp-Massey linear feedback shift register synthesis algorithm. *IBM Journal of Research and Development*, 20:204–212, 1976. 15

[26] Y. K. Han, J.-H. Chang, and K. Yang. On the k-error linear complexity of p^m-periodic binary sequences. *IEEE Trans. Inform. Theory*, 53(6):2297–2304, 2007. 26, 37

[27] M. Hell and T. Johansson. Breaking the f-fcsr-h stream cipher in real time. In J. Pieprzyk, editor, *ASIACRYPT*, volume 5350 of *LNCS*, pages 557–569. Springer, 2008. 85

[28] Y.-P. Hong, Y.-C. Eun, J.-H. Kim, and H.-Y. Song. Linear complexity of sequences over arbitrary symbols and constructions of sequences over $\mathbb{F}_{p^k}$ whose characteristic polynomial is over $\mathbb{F}_p$. In *Intl. Symp. Inform. Theory.*, page 468, 2002. 81

[29] H. Hu, , G. Gong, and Z. Dai. New results on periodic sequences with large k-error linear complexity. *Int. Symp. Inform. Theory*, pages 2409–2413, 2008. 16, 82

[30] H. Hu and D. Feng. On the 2-adic complexity and the k-error 2-adic complexity of periodic binary sequences. In T. Helleseth, D. V. Sarwate, H.-Y. Song, and K. Yang, editors, *SETA 2004*, volume 3486 of *LNCS*, pages 185–196. Springer, 2004. 17, 31

[31] H. Hu, L. Hu, and D. Feng. On the expected value of the joint 2-adic complexity of periodic binary multisequences. In G. Gong, T. Helleseth, H.-Y. Song, and K. Yang, editors, *SETA 2006*, volume 4086 of *LNCS*, pages 199–208. Springer, 2006. 17, 34

[32] S. Jiang, Z. Dai, and K. Imamura. Linear complexity of a sequence obtained from a periodic sequence by either substituting, insertion or deleting k symbols within one period. *IEEE Trans. Inform. Theory*, 46(3):1328–1331, 2000. 18, 19, 23

[33] T. Kaida, S. Uehara, and K. Imamura. A new algorithm for the k-error linear complexity of sequences over $GF(p^m)$ with period p^n. In C. Ding, T. Helleseth, and H. Niederreiter, editors, *SETA 1998*, pages 284–296. Springer, 1999. 37

[34] R. Kavuluru. 2^n-periodic binary sequences with fixed 2-error or 3-error linear complexity. In S. Golomb, M. Parker, A. Pott, and A. Winterhof, editors, *SETA 2008*, volume 5203 of *LNCS*, pages 252–265. Springer, 2008. 57

[35] R. Kavuluru. Characterization of 2^n-periodic binary sequences with fixed 2-error or 3-error linear complexity. *Designs, Codes, and Cryptography*, 53(2):75–97, 2009. 57

[36] R. Kavuluru and A. Klapper. Counting functions for the k-error linear complexity of 2^n-periodic binary sequences. to appear in the LNCS proceedings of the 15th annual workshop on *Selected Areas in Cryptography (SAC 2008)*. 45

[37] R. Kavuluru and A. Klapper. On the k-operation linear complexity of periodic sequences. In K. Srinathan, C. P. Rangan, and M. Yung, editors, *INDOCRYPT 2007*, volume 4859 of *LNCS*, pages 322–330. Springer, 2007. 19

[38] R. Kavuluru and A. Klapper. Lower bounds on error complexity measures for periodic LFSR and FCSR sequences. *Cryptography and Communications: Discrete Structures, Boolean Functions, and Sequences*, 1:95–116, 2009. 19

[39] A. Klapper. The asymptotic behavior of N-adic complexity. *Advances in Mathematics of Communications*, 1(3):307–319, 2007. 11

[40] A. Klapper. Expected π-adic security measures of sequences. In S. Golomb, M. Parker, A. Pott, and A. Winterhof, editors, *SETA 2008*, volume 5203 of *LNCS*, pages 219–229. Springer, 2008. 81

[41] A. Klapper and M. Goresky. Feedback shift registers, combiners with memory, and 2-adic span. *J. Cryptology*, 10:111–147, 1997. 5, 10

[42] A. Klapper and J. Xu. Algebraic feedback shift registers. *Theoretical Computer Science*, 226(1-2):61–92, 1999. 81

[43] K. Kurosawa, F. Sato, T. Sakata, and W. Kishimoto. A relationship between linear complexity and k-error linear complexity. *IEEE Trans. Inform. Theory*, 46(2):694–698, 2000. 25, 64

[44] A. Lauder and K. Patterson. Computing the error linear complexity spectrum of a binary sequence of period 2^n. *IEEE Trans. Inform. Theory*, 49(1):273–280, 2003. 37, 38, 56

[45] R. Lidl and H. Niederreiter. *Finite Fields*. Cambridge University Press, 1997. 4

[46] K. Limniotis, N. Kolokotronis, and N. Kalouptsidis. On the nonlinear complexity and Lempel-Ziv complexity of finite length sequences. *IEEE Trans. Inform. Theory*, 53(11):4293–4302, 2007. 82

[47] G. Marsaglia. The mathematics of random number generators. In S. A. Burr, editor, *The Unreasonable Effectiveness of Number Theory*, volume 42 of *Proc. Symp. Appl. Math.*, pages 73–92. Amer. Math. Soc., 1992. 5

[48] J. L. Massey. Shift register synthesis and BCH decoding. *IEEE Trans. Inform. Theory*, 15(1):122–127, 1969. 8

[49] J. L. Massey and T. Schaub. Linear complexity in coding theory. In G. D. Cohen and P. Godlewski, editors, *Coding Theory and Applications*, volume 311 of *LNCS*, pages 19–32, 1988. 16

[50] R. McEliece. *Finite Fields For Computer Scientists and Engineers*. Kluwer Academic Publishers, 1989. 4

[51] W. Meidl. Extended Games-Chan algorithm for the 2-adic complexity of FCSR sequences. *Theoretical Computer Science*, 290(3):2045–2051, 2003. 37

[52] W. Meidl. How many bits have to be changed to decrease the linear complexity? *Designs, Codes, and Cryptography*, 33(2):109–122, 2004. 26, 37

[53] W. Meidl. On the stability of 2^n-periodic binary sequences. *IEEE Trans. Inform. Theory*, 51(3):1151–1155, 2005. 37, 44, 45, 46, 65

[54] W. Meidl. Reducing the calculation of the linear complexity of $u2^v$-periodic binary sequences to Games-Chan algorithm. *Designs, Codes, and Cryptography*, 46(1):57–65, 2008. 37

[55] W. Meidl and H. Niederreiter. Counting functions and expected values for the k-error linear complexity. *Finite Fields and Applications*, 8:142–154, 2002. 15

[56] W. Meidl and H. Niederreiter. Linear complexity, k-error linear complexity, and the discrete fourier transform. *J. Complexity*, 18(1):87–103, 2002. 16, 37

[57] W. Meidl and H. Niederreiter. On the expected value of linear complexity and the k-error linear complexity of periodic sequences. *IEEE Trans. Inform. Theory*, 48(11):2817–2825, 2002. 16

[58] W. Meidl and H. Niederreiter. The expected value of the joint linear complexity of periodic multisequences. *J. Complexity*, 19(1):61–72, 2003. 12, 13, 16

[59] W. Meidl and H. Niederreiter. Periodic sequences with maximal linear complexity and large k-error linear complexity. *Appl. Algebra Eng. Commun. Comput.*, 14(4):273–286, 2003. 16

[60] W. Meidl, H. Niederreiter, and A. Venkateswarlu. Error linear complexity measures for multisequences. *J. Complexity*, 23(2):169–192, 2007. 13, 16

[61] W. Meidl and A. Venkateswarlu. Remarks on the k-error linear complexity of p^n-periodic sequences. *Designs, Codes, and Cryptography*, 42(2):181–193, 2007. 37

[62] W. Meier and O. Staffelbach. Fast correlation attacks on certain stream ciphers. *J. Cryptology*, 1(3):159–176, 1989. 84

[63] H. Niederreiter. A combinatorial approach to probabilistic results on the linear complexity profile of random sequences. *J. Cryptology*, 2(2):105–112, 1990. 11, 16

[64] H. Niederreiter. Linear complexity and related complexity measures for sequences. In T. Johansson and S. Maitra, editors, *INDOCRYPT*, volume 2904 of *LNCS*, pages 1–17. Springer, 2003. 16

[65] H. Niederreiter. Periodic sequences with large k-error linear complexity. *IEEE Trans. Inform. Theory*, 49(2):501–505, 2003. 16

[66] H. Niederreiter and I. Shparlinski. Recent advances in the theory of nonlinear pseudorandom number generators. In K. T. Fang, F. J. Hickernell, and H. Niederreiter, editors, *MCQMC 2000*, pages 86–102. Springer-Verlag, 2002. 8

[67] H. Niederreiter and I. Shparlinski. Periodic sequences with maximal linear complexity and almost maximal k-error linear complexity. In K. G. Paterson, editor, *IMA Int. Conf.*, volume 2898 of *LNCS*, pages 183–189. Springer, 2003. 16

[68] H. Niederreiter and A. Venkateswarlu. Periodic multisequences with large error linear complexity. *Designs, Codes, and Cryptography*, 49(1-3):33–45, 2008. 16

[69] H. Niederreiter and L.-P. Wang. Proof of a conjecture on the joint linear complexity profile of multisequences. In S. Maitra, C. E. V. Madhavan, and R. Venkatesan, editors, *INDOCRYPT 2005*, volume 3797 of *LNCS*, pages 13–22. Springer, 2005. 12

[70] K. Patterson. Perfect maps. *IEEE Trans. Inform. Theory*, 40(3):743–753, 1994. 38

[71] P. Rizomiliotis. Constructing periodic binary sequences with maximum nonlinear span. *IEEE Trans. Inform. Theory*, 52(9):4257–4261, 2006. 82

[72] P. Rizomiliotis. Remarks on the new attack on the filter generator and the role of high order complexity. In S. D. Galbraith, editor, *IMA Int. Conf.*, volume 4887 of *LNCS*, pages 204–219. Springer, 2007. 82

[73] P. Rizomiliotis and N. Kalouptsidis. Results on the nonlinear span of binary sequences. *IEEE Trans. Inform. Theory*, 51(4):1555–1563, 2005. 82

[74] P. Rizomiliotis, N. Kolokotronis, and N. Kalouptsidis. On the quadratic span of binary sequences. *IEEE Trans. Inform. Theory*, 51(5):1840–1848, 2005. 82

[75] R. A. Rueppel. *New Approaches to Stream Ciphers*. PhD thesis, Swiss Federal Institute of Research, 1984. 5

[76] R. A. Rueppel. *Analysis and Design of Stream Ciphers*. Springer, 1986. 10, 39, 84

[77] A. Sălăgean. On the computation of the linear complexity and the k-error linear complexity of binary sequences with period a power of two. *IEEE Trans. Inform. Theory*, 51(3):1145–1150, 2005. 38

[78] V. M. Sidel'nikov. Some k-valued pseudorandom sequences and nearly equidistant codes. *Probl. Peredachi Inf.*, 5(1):16–22, 1969. 38

[79] T. Siegenthaler. Decrypting a class of stream ciphers using ciphertext only. *IEEE Trans. on Computers*, C-34(1):81–85, 1985. 84

[80] M. Stamp and C. F. Martin. An algorithm for the k-error linear complexity of binary sequences with period 2^n. *IEEE Trans. Inform. Theory*, 39(4):1398–1401, 1993. 11, 37

[81] M. Su and L. Chen. The properties of the 1-error linear complexity of p^n-periodic sequences over $\mathbb{F}_p$. *Int. Symp. Inform. Theory*, pages 1998–2002, 2006. 57

[82] S. Uehara and K. Imamura. Linear complexity of periodic sequences obtained from $\mathbb{F}_q$ sequences with period $q^n - 1$ by one-symbol insertion. *IEICE Trans. Fundamentals*, E79-A:1739–1740, 1996. 18

[83] S. Uehara and K. Imamura. Linear complexity of periodic sequences obtained from a sequence over $\mathbb{F}_p$ with period $p^n - 1$ by one-symbol deletion. *IEICE Trans. Fundamentals*, E80-A:1164–1166, 1997. 18

[84] L.-P. Wang. A lattice-based minimal partial realization algorithm. In S. Golomb, M. Parker, A. Pott, and A. Winterhof, editors, *SETA 2008*, volume 5203 of *LNCS*, pages 278–289. Springer, 2008. 81

[85] L.-P. Wang and Niederreiter. Enumeration results on the joint linear complexity of multisequences. *Finite Fields and Applications*, 12:613–637, 2006. 16

[86] L.-P. Wang and Niederreiter. The asymptotic behavior of the joint linear complexity profile of multisequences. *Monatshefte für mathematik*, 150(2):141–155, 2007. 16

[87] L.-P. Wang, Y.-F. Zhu, and D.-Y. Pei. On the lattice basis reduction multisequence synthesis algorithm. *IEEE Trans. Inform. Theory*, 50(11):2905–2910, 2004. 81

[88] G. Xiao and S. Wei. Fast algorithms for determining the linear complexity of periodic sequences. In A. Menezes and P. Sarkar, editors, *INDOCRYPT 2002*, volume 2551 of *LNCS*, pages 12–21. Springer, 2002. 37

[89] G. Xiao, S. Wei, K.-Y. Lam, and K. Imamura. A fast algorithm for determining the linear complexity of a sequence with period p^n over $\mathbb{F}_q$. *IEEE Trans. Inform. Theory*, 46(1):2203–2206, 2000. 37

[90] J. Xu. *Stream Cipher Analysis Based on FCSRs*. PhD thesis, University of Kentucky, 2000. 10

Vita

Name Venkata Naga Rama Kanth Kavuluru

Place of Birth Vijayawada, India

Education

- M.S, Computer Science, Western Kentucky University, 2004.
- B.Tech, Computer Science and Eng., Jawaharlal Nehru Tech. University, 2002.

Work Experience

- 2004–2009, Teaching and Research Assistant, University of Kentucky.
- 2003–2004, Research Assistant, Western Kentucky University.
- 2003–2006, Summer Teacher, Suffield Academy.

Honors

- Awarded Harrision D. Brailsford graduate fellowship for 2008–2009
- Awarded Verizon graduate fellowship for 2008–2009.
- Awarded Kentucky Graduate Scholarship for 2004–2009

Journal Articles

- "Numerical Upper Bounds on Rope Lengths of Large Physical Knots," by Y. Diao, C. Ernst, R. Kavuluru, and U. Ziegler, *Journal of Physics A: Math. Gen.* Vol. **39** (2006), pp. 4829–4843.
- "Lower Bounds on Error Complexity Measures for Periodic LFSR and FCSR Sequences," by R. Kavuluru and A. Klapper, *Cryptography and Communications* Vol. **1** (2009), pp. 95–116.
- "Characterization of 2^n-Periodic Binary Sequences with Fixed 2-Error or 3-Error Linear Complexity," by R. Kavuluru, *Designs, Codes, and Cryptography* Vol. **53**(2) (2009), pp. 75–97.

Conference Proceedings

- "On the k-operation Linear Complexity of Periodic Sequences," by R. Kavuluru and A. Klapper, *Progress in Cryptology* INDOCRYPT 2007, *LNCS* **4859** (2007), pp. 322–330.

- "2^n-Periodic Sequences with Fixed k-error Linear Complexity for $k = 2$ or 3," by R. Kavuluru, 5th international conference on *SEquences and Their Applications (SETA 2008)*, *LNCS* **5203** (2008), pp. 252–265.

- "Counting Functions for the k-error Linear Complexity of 2^n-Periodic Binary Sequences," by R. Kavuluru and A. Klapper, to appear in the LNCS proceedings of 15th annual workshop on *Selected Areas in Cryptography*.

Talks

- Presented papers at conferences INDOCRYPT 2007, SAC 2008, and SETA 2008.

- "Generation Primitives and Complexity Measures for Pseudorandom Sequences," R. Kavuluru, Spring 2008, Western Kentucky University.

- "Enumeration Results on the Linear Complexity of Sequences," R. Kavuluru, Spring 2009, Graduate Student Combinatorics Conference, University of Kentucky.

Academic Service

- Reviewed papers for IEEE Symp. on Information Theory 2006, INDOCRYPT 2009, Cryptography and Communications, and Applicable Algebra in Engineering, Communication, and Computing.

- Maintained the online paper submission/review system and provided audio-visual support for the 5th international conference on sequences and their applications, SETA 2008.

- Volunteered as a microteach leader for the new TA orientation at UKY in 2007.

Professional memberships

International Association for Cryptologic Research