RANDOMNESS PROPERTIES OF SOME VECTOR SEQUENCES GENERATED BY MULTIVARIATE POLYNOMIAL ITERATIONS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

PINAR GÜRKAN BALIKÇIOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
CRYPTOGRAPHY

FEBRUARY 2016

Approval of the thesis:

RANDOMNESS PROPERTIES OF SOME VECTOR SEQUENCES GENERATED BY MULTIVARIATE POLYNOMIAL ITERATIONS

submitted by **PINAR GÜRKAN BALIKÇIOĞLU** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Department of Cryptography, Middle East Technical University** by,

Prof. Dr. Bülent Karasözen
Director, Graduate School of **Applied Mathematics** ________________

Prof. Dr. Ferruh Özbudak
Head of Department, **Cryptography** ________________

Assoc. Prof. Dr. Melek Diker Yücel
Supervisor, **Electrical And Electronics Engineering Department, METU** ________________

Examining Committee Members:

Prof. Dr. Ersan Akyıldız
Mathematics, METU ________________

Assoc. Prof. Dr. Melek Diker Yücel
Electrical And Electronics Engineering Department, METU ________________

Prof. Dr. Ferruh Özbudak
Mathematics, METU ________________

Assist. Prof. Dr. Oğuz Yayla
Mathematics, Hacettepe University ________________

Assist. Prof. Dr. Çetin Ürtiş
Mathematics, TOBB ETÜ ________________

Date: ________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: PINAR GÜRKAN BALIKÇIOĞLU

Signature :

ABSTRACT

RANDOMNESS PROPERTIES OF SOME VECTOR SEQUENCES GENERATED BY MULTIVARIATE POLYNOMIAL ITERATIONS

Gürkan Balıkçıoğlu, Pınar
Ph.D., Department of Cryptography
Supervisor : Assoc. Prof. Dr. Melek Diker Yücel

February 2016, 105 pages

We examine the randomness properties of the sequences generated by the multivariate polynomial iterations method proposed by Ostafe and Shparlinski, by using the six different choices of polynomials given by the same authors. Our analysis is based on two approaches: distributions of the periods and linear complexities of the produced vector sequences. We define the efficiency parameters, PE for "period efficiency" and LCE for "linear complexity efficiency", so that the actual values of the period and linear complexity of a sequence can be easily compared with those of the ideal cases. For each polynomial choice, in order to obtain the period distribution of the generated vector sequences, we perform an exhaustive search for prime field sizes up to 13; and observe that the probability of attaining a maximum-period sequence is extremely low. Linear complexities of the sequences are also computed exhaustively for prime field sizes up to 13 and the multivariate polynomial iterations with the proposed polynomial choices are observed to generate sequences with having high linear complexities quite seldomly. We then concentrate on the largest period sequences produced by each choice, and investigate the linear complexity of those sequences for a given polynomial choice, at a specific field size p and number of polynomials m. We observe that an increase of p or m does not bring any improvement on the randomness of the generated sequences. Finally, we analyze the linear complexity of Ostafe's sequences by fixing the period but leaving the choice of m and other initial values random, as in real life. Although computational constraints limit our exhaustive search results in the first part to relatively small values of p and m; the last part of our study lets us use higher values of p and m, to justify the projection that Ostafe's method with the proposed polynomial choices is not a promising way of implementing pseudo-random number generators.

Keywords: multivariate polynomial iterations, pseudo-random vector sequences, triangular polynomial systems, period distribution, linear complexity

ÖZ

ÇOK DEĞİŞKENLİ POLİNOM TEKRARLAMALARI İLE ÜRETİLEN BAZI VEKTÖR DİZİLERİNİN RASSALLIK ÖZELLİKLERİ

Gürkan Balıkçıoğlu, Pınar
Doktora, Kriptografi Bölümü
Tez Yöneticisi : Doç. Dr. Melek Diker Yücel

Ocak 2016, 105 sayfa

Ostafe ve Shparlinski tarafından önerilen çok değişkenli polinom iterasyonları ile üretilen dizilerin, aynı yazarlar tarafından önerilen altı polinom seçeneği için rassallık özelliklerini araştırdık. Analizimiz iki yaklaşımı temel almaktadır: üretilen vektör dizilerinin periyot ve doğrusal karmaşıklık dağılımı. Elde ettiğimiz değerleri, bu yaklaşımların ideal durumları ile karşılaştırabilmek amacıyla, periyot yeterliliği için PE ve doğrusal karmaşıklık yeterliği için LCE olmak üzere, yeterlik parametreleri tanımladık. Üretilen vektör dizilerinin, her bir polinom seçeneği için periyot dağılımını elde edebilmek amacıyla, büyüklüğü 13'e kadar olan asal alanlar için tüm olası durumlar üzerinden araştırma yaptık ve maksimum uzunluklu dizilere erişme olasılığının oldukça düşük olduğunu gözlemledik. Ayrıca büyüklüğü 13'e kadar olan asal alanlar için tüm olası durumlar üzerinden dizilerin doğrusal karmaşıklıklarını araştırdık ve çok değişkenli polinom iterasyonları metoduyla birlikte önerilen polinom seçimlerinin, oldukça seyrek durumlarda yüksek doğrusal karmaşıklığa sahip diziler üretebildiğini gözlemledik. Ardından, her seçim tarafından üretilen en yüksek periyotlu dizilere yoğunlaştık ve bu dizilerin verilen bir seçim için, belirli bir alan büyüklüğü (p) ve polinom sayısındaki (m) doğrusal karmaşıklığını inceledik. p ve m'yi arttırmanın üretilen dizilerin rassallığına dair herhangi bir iyileşme sağlamadığını gözlemledik. Son olarak, gerçek hayattaki gibi periyotları sabit tutup, m ve diğer başlangıç değerlerini rassal olarak alıp; Ostafe'nin dizilerinin doğrusal karmaşıklığını analiz ettik. İlk bölümdeki kapsamlı araştırmamızı hesaplama zorlukları nedeniyle nispeten küçük p ve m değerleri ile sınırlamamıza rağmen; son bölümdeki çalışmamız daha büyük p ve m değerleri kullanmamıza izin vererek, önerilen polinom seçenekleri ile kullanıldığında Ostafe'nin yönteminin rassal sayı üreteci olarak gerçeklenemeyeceği konusundaki çıkarımımızı desteklemektedir.

Anahtar Kelimeler : çok değişkenli polinom iterasyonları, sözde rassal vektör dizileri, periyot dağılımı, doğrusal karmaşıklık

To My Family

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to my thesis supervisor, Assoc. Prof. Dr. Melek Diker Yücel, for her patience, motivation, and immense knowledge throughout not only my thesis work but also my life in general. This Ph.D study would not have been completed without her patience and guidance.

Besides my supervisor, I would like to acknowledge Prof. Dr. Ersan Akyıldız and Prof. Dr. Ferruh Özbudak, for their insightful comments which steered me to extend my research from various perspectives.

I would also like to give my special thanks to my husband Servet Balıkçıoğlu for his trust in me. His endless patience has enlightened my path and his love has been always my shelter.

I am grateful to my parents Hikmet Gürkan and İhsan Gürkan for their love and unfailing support during my whole life.

Finally, I must express my very profound gratitude to all my friends for always trusting and supporting me.

TABLE OF CONTENTS

ABSTRACT vii

ÖZ ix

ACKNOWLEDGMENTS xiii

TABLE OF CONTENTS xv

LIST OF FIGURES xix

LIST OF TABLES xxi

CHAPTERS

1 INTRODUCTION 1

1.1 Pseudo Random Noise Generators 1

1.1.1 Linear Recurrence Generator Modulo 2 2

1.1.2 k-th Order Linear Recurrence Generator 2

1.1.3 Non-Linear Congruential Generator 3

1.2 Ostafe and Shparlinski's Method 3

1.3 Testing Randomness 6

1.4 Aim and Skeleton of the Thesis 6

2 EXHAUSTIVE PERIOD ANALYSIS 9

2.1 Introduction 9

2.2 Set of Parameters 9

2.3 Period Distributions 11

2.4 Factors of Periods 16

2.5 Conclusion 17

3 EXHAUSTIVE LINEAR COMPLEXITY ANALYSIS 19

3.1 Introduction 19

3.2 Linear Complexity 19

3.3 Measuring Randomness in Terms of the Linear Complexity . 20

3.4 Linear Complexity Distributions 21

3.5 Minimum, Average and Maximum LCE versus T_v of the Sequences Generated by the Five Polynomial Choices 25

3.6 Linear Complexity Efficiency of Sequences with Period Efficiency, PE ≥ 0.5 25

3.7 Sequences with High Linear Complexity Efficiency, LCE $\geq$ 0.95 30

3.7.1 Percentages 30

3.7.2 Corresponding Vector Periods 31

3.8 Conclusion 31

4 LINEAR COMPLEXITY ANALYSIS OF THE LARGEST PERIOD SEQUENCES 33

4.1 Introduction 33

4.2 LCE's Obtained at the Largest T_v of the Corresponding Choices versus the Field Size p 33

4.3 LCE's Obtained at the Largest T_v of the Corresponding Choices versus the Vector Size m for $p = 3$ 34

4.4 Linear Complexity of Maximum-Period Sequences 37

4.4.1 Generating Maximal Period Sequences with Choice 4 37

4.4.2 Linear Complexity Efficiency Values of Choice 4 Sequences 38

4.4.3 Example Sequences Generated by Choice 4 with Poor Randomness 38

4.5 Conclusion 41

5 LINEAR COMPLEXITY ANALYSIS OF SEQUENCES GENERATED WITH RANDOM INITIAL VALUES 43

5.1 Introduction 43

5.2 Randomness Comparison with MATLAB's **randi(.)** Sequence 43

5.3 Effect of the Field Size p for Similar Period Sequences 45

5.4 Effect of the Number of Polynomials m for Similar Period Sequences 47

5.5 Effect of p and m for Variable Period Sequences 48

5.6 Conclusion 49

6 CONCLUSION 53

REFERENCES 55

APPENDICES

A COMPUTATION TIMES OF THE EXHAUSTIVE SEARCHES FOR FIVE CHOICES 57

B CHOICE 1 SEQUENCES OVER $\mathbb{F}_2$ 59

C EXHAUSTIVE PERIOD ANALYSIS 61

D EXHAUSTIVE LINEAR COMPLEXITY ANALYSIS 67

E LINEAR COMPLEXITY ANALYSIS OF THE LARGEST PERIOD SEQUENCES 77

F LINEAR COMPLEXITIES OF SEQUENCES GENERATED WITH RANDOM INITIAL VALUES 83

G PERCENTAGES OF HIGH-LCE SEQUENCES GENERATED WITH RANDOM INITIAL VALUES AND CORRESPONDING PE'S . . . 87

H LCE VALUES OF THE SEQUENCES GENERATED BY CHOICE 3a WITH RANDOM INITIAL VALUES FOR $T = 1000$ 91

I MATLAB IMPLEMENTATIONS OF ALGORITHMS IN THIS WORK 95

I.1 MATLAB Implementation of Choice 1 95

I.2 MATLAB Implementation of Choice 2 96

I.3 MATLAB Implementation of Choice 3a 96

I.4 MATLAB Implementation of Choice 3b 97

I.5 MATLAB Implementation of Choice 3c 97

I.6 MATLAB Implementation of Choice 4 98

I.7 MATLAB Implementation of the Berlekamp-Massey Algorithm 99

J LINEAR COMPLEXITY PROFILES 101

CURRICULUM VITAE 105

LIST OF FIGURES

Figure 2.1 Distribution of the period efficiency, PE (the period T_v of the vector sequence divided by the maximum period p^m), for Choice 1, 2, 3a, 3b, 3c and 4 and the field size $p = 3$, where the number of polynomials are: (a) $m = 2$, (b) $m = 3$, (c) $m = 4$, (d) $m = 5$ 12

Figure 2.2 Distribution of the period efficiency, PE (the period T_v of the vector sequence divided by the maximum period p^m), for Choice 1, 2, 3a, 3b, 3c and 4 and the field sizes: (a) $p = 5$, (b) $p = 7$, (c) $p = 11$, (d) $p = 13$ and the number of polynomials $m = 2$ 13

Figure 2.3 Distribution of the period efficiency, PE (the period T_v of the vector sequence divided by the maximum period p^m), for Choice 1, 2, 3a, 3b, 3c and 4; where the field size p and the number of polynomials m are given as (a) $p = 5, m = 2$, (b) $p = 5, m = 3$, (c) $p = 7, m = 2$, (d) $p = 7, m = 3$ . . 14

Figure 3.1 Distribution of the linear complexity efficiency, LCE (ratio of the computed linear complexity L of the sequence to its length T), for Choice 1, 2, 3a, 3b and 3c (respectively dark blue, red, green, purple, light blue and orange) and the field size $p = 3$, where the number of polynomials are: (a) $m = 2$, (b) $m = 3$, (c) $m = 4$, (d) $m = 5$ 22

Figure 3.2 Distribution of the linear complexity efficiency, LCE (ratio of the computed linear complexity L of the sequence to its length T), for Choice 1, 2, 3a, 3b and 3c and the field sizes: (a) $p = 5$, (b) $p = 7$, (c) $p = 11$, (d) $p = 13$ and the number of polynomials $m = 2$ 23

Figure 3.3 Distribution of the linear complexity efficiency, LCE (ratio of the computed linear complexity L of the sequence to its length T), for Choice 1, 2, 3a, 3b and 3c; where the field size p and the number of polynomials m are given as (a) $p = 5, m = 2$, (b) $p = 5, m = 3$, (c) $p = 7, m = 2$, (d) $p = 7, m = 3$ 24

Figure 3.4 LCE values versus vector period T_v of sequences for $p = 5, m = 3$ generated by (a) Choice 1, (b) Choice 2, (c) Choice 3a polynomials 26

Figure 3.5 LCE values versus vector period T_v of sequences for $p = 7, m = 3$ generated by (a) Choice 1, (b) Choice 2, (c) Choice 3a polynomials 27

Figure 3.6 LCE values versus vector period T_v of sequences for $p = 11, m = 2$ generated by (a) Choice 1, (b) Choice 2, (c) Choice 3a polynomials 28

Figure 3.7 LCE values versus vector period T_v of sequences for $p = 13, m = 2$ generated by (a) Choice 1, (b) Choice 2, (c) Choice 3a polynomials 29

Figure 4.1 Variation of the: (a) minimum, (b) average and (c) maximum LCE's obtained at the largest T_v of the corresponding choice versus p, at $m = 2$. 35

Figure 4.2 Minimum, average and maximum LCE obtained at the largest T_v of: (a) Choice 1, (b) Choice 2, (c) Choice 3a, (d) Choice 3b and (e) Choice 3c, versus m for $p = 3$ 36

Figure 5.1 Randomness comparison of Ostafe's and MATLAB's **randi(.)** sequences for (a) $p = 3$, (b) $p = 13$, in terms of the LCE as defined in this work by "linear complexity L divided by the period T of the sequence" (For each T, the average LCE is computed over 100 sequences.) 44

Figure 5.2 Percentage of sequences with LCE ≥ 0.95 for (a) $p = 3$, (b) $p = 13$, using three choices 45

Figure 5.3 LCE values for a sequence of length (a) $T = 500$, (b) $T = 800$, (c) $T = 1000$, in variable field sizes, $3 \leq p \leq 31$ using the three choices given in Table 1.1 46

Figure 5.4 Variation of the average LCE's of the sequences generated by random initialization of the corresponding choice versus p at: (a) $m = 2$, (b) $m = 3$, (c) $m = 4$ 50

Figure 5.5 Variation of the minimum LCE's of the sequences generated by random initialization of the corresponding choice versus p at: (a) $m = 2$, (b) $m = 3$, (c) $m = 4$ 51

Figure F.1 Randomness comparison of the three choices with the random sequences produced by MATLAB's **randi(.)** where (a) $p = 5$, (b) $p = 7$, (c) $p = 11$, in terms of the LCE as defined in this work by "linear complexity L divided by the period T of the sequence" (For each T, average LCE is computed over 100 sequences.) 84

Figure F.2 Percentage of sequences with LCE ≥ 0.95 for (a) $p = 5$, (b) $p = 7$, (c) $p = 11$ using three choices 85

Figure J.1 Linear complexity profile of three sequences generated by Choice 1 with high (LCE ≥ 0.95) where the field size p and the scalar period T are given as (a) $p = 5, T = 100$, (b) $p = 5, T = 100$, (c) $p = 7, T = 500$, (d) $p = 7, T = 500$ 102

Figure J.2 Linear complexity profile of three sequences with LCE < 0.7 generated by Choice 1 of the scalar period $T = 100$ where the field size p is given as (a) $p = 5$, (b) $p = 7$ 103

LIST OF TABLES

Table 1.1 Polynomial choices 5

Table 2.1 Size of parameter sets 10

Table 2.2 Number of possible vector sequences 10

Table 2.3 Number of possible vector sequences for the ten cases 11

Table 2.4 Weighted average of PE values for ten cases 15

Table 2.5 Percentage of Choice 2 sequences with maximum-period p^m 15

Table 2.6 Percentage of sequences with period efficiency ≤ 0.5 16

Table 2.7 The highest vector periods of the corresponding polynomial choices 17

Table 3.1 Weighted average of LCE values for ten cases 21

Table 3.2 Average linear complexity efficiency of the sequences with period efficiency ≥ 0.5, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m and h_m 30

Table 3.3 Percentage of sequences with linear complexity efficiency LCE $\geq$ 0.95, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m and h_m 31

Table 3.4 Vector periods of sequences with linear complexity efficiency LCE $\geq$ 0.95, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m and h_m 32

Table 4.1 Linear complexity efficiency values of the maximum-period sequences 39

Table 4.2 Linear complexity efficiency values of the maximum-period sequences for $p = 2$ 40

Table 4.3 An example of the sequences generated by Choice 4 40

Table 5.1 LCE values of the sequences generated by Choice 1 for $p = 3$, $T =$ 1000, $m = 9$ 47

Table 5.2 LCE values of the sequences generated by Choice 2 for $p = 3$, $T =$ 1000, $m = 12$ 48

Table 5.3 LCE values of the sequences generated by Choice 3a for $p = 3$, $T = 1000, m = 19$ 48

Table 5.4 LCE values of the sequences generated by Choice 3a for $p = 3$, $T = 1000$ 48

Table A.1 Computation times (in seconds) of the exhaustive PE and LCE searches for five choices 58

Table C.1 Percentage of sequences with vector period $T_v = 1$ 62

Table C.2 The most common vector periods of the corresponding polynomial choices 62

Table C.3 Vector periods obtained with two polynomials ($m = 2$) 63

Table C.4 Vector periods obtained with three polynomials ($m = 3$) 64

Table C.5 Vector periods obtained with four polynomials ($m = 4$) 65

Table C.6 Vector periods obtained with five polynomials ($m = 5$) 65

Table D.1 LCE values versus vector period T_v of sequences generated by Choice 1 values for $p = 5, m = 3$ 68

Table D.2 LCE values versus vector period T_v of sequences generated by Choice 2 values for $p = 5, m = 3$ 68

Table D.3 LCE values versus vector period T_v of sequences generated by Choice 3a values for $p = 5, m = 3$ 69

Table D.4 LCE values versus vector period T_v of sequences generated by Choice 1 values for $p = 7, m = 3$ 69

Table D.5 LCE values versus vector period T_v of sequences generated by Choice 2 values for $p = 7, m = 3$ 70

Table D.6 LCE values versus vector period T_v of sequences generated by Choice 3a values for $p = 7, m = 3$ 71

Table D.7 LCE values versus vector period T_v of sequences generated by Choice 1 values for $p = 11, m = 2$ 71

Table D.8 LCE values versus vector period T_v of sequences generated by Choice 2 values for $p = 11, m = 2$ 72

Table D.9 LCE values versus vector period T_v of sequences generated by Choice 3a values for $p = 11, m = 2$ 72

Table D.10LCE values versus vector period T_v of sequences generated by Choice 1 values for $p = 13, m = 2$ 73

Table D.11LCE values versus vector period T_v of sequences generated by Choice 2 values for $p = 13, m = 2$ 74

Table D.12LCE values versus vector period T_v of sequences generated by Choice 3a values for $p = 13, m = 2$ 75

Table D.13Minimum linear complexity efficiency of the sequences with period efficiency PE ≥ 0.5, found over all possible values of **X**, **G**, **H**, a_i, g_m and h_m 75

Table D.14Maximum linear complexity efficiency of the sequences with period efficiency PE ≥ 0.5, found over all possible values of **X**, **G**, **H**, a_i, g_m and h_m 76

Table E.1 LCE values obtained at the largest T_v of Choice 1 at $m = 2$ 78

Table E.2 LCE values obtained at the largest T_v of Choice 2 at $m = 2$ 78

Table E.3 LCE values obtained at the largest T_v of Choice 3a at $m = 2$ 78

Table E.4 LCE values obtained at the largest T_v of Choice 3b at $m = 2$ 78

Table E.5 LCE values obtained at the largest T_v of Choice 3c at $m = 2$ 79

Table E.6 LCE values obtained at the largest T_v of Choice 1 versus m for $p = 3$ 79

Table E.7 LCE values obtained at the largest T_v of Choice 2 versus m for $p = 3$ 79

Table E.8 LCE values obtained at the largest T_v of Choice 3a versus m for $p = 3$ 80

Table E.9 LCE values obtained at the largest T_v of Choice 3b versus m for $p = 3$ 80

Table E.10LCE values obtained at the largest T_v of Choice 3c versus m for $p = 3$ 81

Table G.1 Percentages of high-LCE sequences (LCE ≥ 0.95) with $T_v \geq p - 1$ and corresponding maximum PE's generated by random initialization at $m = 2$ 88

Table G.2 Percentages of high-LCE sequences (LCE ≥ 0.95) with $T_v \geq p - 1$ and corresponding maximum PE's generated random initialization at $m = 3$ 88

Table G.3 Percentages of high-LCE sequences (LCE ≥ 0.95) with $T_v \geq p - 1$ and corresponding maximum PE's generated by random initialization at $m = 4$ 89

Table H.1 LCE values of the randomly initiated sequences generated by Choice 3a for $p = 5, T = 1000$ 91

Table H.2 LCE values of the randomly initiated sequences generated by Choice 3a for $p = 7, T = 1000$ 92

Table H.3 LCE values of the randomly initiated sequences generated by Choice 3a for $p = 11, T = 1000$ 92

Table H.4 LCE values of the randomly initiated sequences generated by Choice 3a for $p = 13, T = 1000$ 92

Table H.5 LCE values of the randomly initiated sequences generated by Choice 3a for $p = 17, T = 1000$ 92

Table H.6 LCE values of the randomly initiated sequences generated by Choice 3a for $p = 19, T = 1000$ 92

Table H.7 LCE values of the randomly initiated sequences generated by Choice 3a for $p = 23, T = 1000$ 92

Table H.8 LCE values of the randomly initiated sequences generated by Choice 3a for $p = 29, T = 1000$ 93

Table H.9 LCE values of the randomly initiated sequences generated by Choice 3a for $p = 31, T = 1000$ 93

CHAPTER 1

INTRODUCTION

In cryptographic applications, random sequences are needed and used frequently. However, obtaining a truly random sequence is very hard and time consuming in real life. Usually, deterministic processes are used to generate pseudo-random sequences as powerful alternatives for truly random sequences. A mathematical algorithm, which generates a pseudo-random sequence starting from a given initial state, is called a *Pseudo-Random Number Generator* (PRNG).

Pseudo-random sequences are produced in a systematic way such that they are statistically indistinguishable from a truly random sequence and their successively generated versions are independent of each other. Sometimes pseudo-random numbers can appear more random than the random numbers obtained from true *Random Number Generator*s (RNGs). Each value in a suitably constructed pseudo-random sequence is produced from the previous value along transformations, which advance extra randomness. Statistical auto-correlations between the input and the output can be eliminated by a series of such transformations. Hence, a PRNG may be faster and generate sequences with better statistical properties than RNGs [29].

1.1 Pseudo Random Noise Generators

Fundamental information on pseudo-random number generation can be found in the book of Knuth [7]. In general, a pseudo-random number generator is based on three generators: a linear recurrence generator modulo 2, a k-th order linear recurrence generator and a non-linear congruential generator [6].

1.1.1 Linear Recurrence Generator Modulo 2

Definition 1.1. *Let $a_1, \ldots, a_{k-1} \in \{0, 1\}$, $a_k = 1$, and each b_j take a value in $\{0, 1\}$. The recurrence relation,*

$$b_i \equiv \sum_{j=1}^{k} a_j b_{i-j} \text{ (mod 2)}$$

generates a sequence $\{b_i\}$ of 0s and 1s. This method is called a linear recurrence generator modulo 2, which is the basis of shift register generators.

1.1.2 k-th Order Linear Recurrence Generator

The formula of k-th order linear recurrence generator is given by following definition.

Definition 1.2. *Let $a_0, \ldots, a_{k-1}$ and c be non-negative integers with $a_0 \neq 0$ and M be a positive integer. Then, x_{i+k} can be computed as*

$$x_{i+k} \equiv \sum_{j=1}^{k} a_{k-j} x_{i+k-j} + c \text{ (mod } M\text{)}, \ \ 0 \leq x_i < M.$$

The first order linear recurrence generator is the linear congruential generator (LCG) invented by Lehmer [11]. However, the output of LCGs are predictable [8]. In [2] and [3], a solution on the predictableness of some LCGs has been proposed. An algorithm that can guess any output of an LCG in its general form is given in [9]. The insecurity of using such generators for cryptographical purposes is also emphasized in [27].

Definition 1.3. *Let M be a large positive integer, a be an integer with $0 < a < M$ and c be an integer with $0 \leq c < M$. Then, an initial value (seed) x_0, $0 \leq x_0 < M$, is selected and a sequence $x_0, x_1, \ldots$ is generated by the recursion*

$$x_{i+1} = a x_i + c \, (mod \, M), \ \ i \geq 0.$$

In this context, M is referred to as the *modulus*, a as the *multiplier* and c as the *increment*. A common distinction is made between the homogeneous case where $c = 0$, also called the *multiplicative congruential method*, and the inhomogeneous case where $c \neq 0$, also called the *mixed congruential method.*

1.1.3 Non-Linear Congruential Generator

Definition 1.4. *Let f be a nonlinear integer-valued function of x_i. Then, a nonlinear congruential generator takes the form*

$$x_{i+1} \equiv f(x_i) \ \textit{(mod M)}; \ \ 0 \leq x_{i+1} < M.$$

An example of a nonlinear congruential generator is the inversive congruential generator suggested by Eichenauer, Grothe, and Lehn [5]. It uses the modular multiplicative inverse (if it exists) to generate the next number in a sequence.

Definition 1.5. *Let a and b be some the positive integers and the prime p be the modulus. The formula of an inversive congruential generator is:*

$$x_{i+1} \equiv (ax_i^{-1} + c) \ \textit{mod} \ p, \ \ \textit{where} \ \ x_0 \neq 0.$$

1.2 Ostafe and Shparlinski's Method

Ostafe and Shparlinski have been inspired by the linear congruential method in their proposals [18]-[25], for generating pseudo-random vector sequences. They study a class of dynamical systems generated by iterations of multivariate polynomials. In [22], this construction is used for designing a new class of hash functions. In [23], linear independence of iterates is studied. The discrepancy of pseudo-random vectors is estimated in [24] and [26]. Joint linear complexity profile of a class of non-linear pseudo-random multi-sequences are studied in [25].

The method of Ostafe and Shparlinski is described as follows: Let p be a prime and $F_1, \ldots, F_m \in \mathbb{F}_p[X_1, \ldots, X_m]$ be m polynomials in m variables over a finite field of p elements. For each $i = 1, \ldots, m$, the k-th iteration of the polynomial F_i is defined by the recurrence relation,

$$f_i^{(k+1)} = F_i(f_1^{(k)}, \ldots, f_m^{(k)}), \ \forall k \ \text{ where } \ f_i^{(0)} = X_i. \tag{1.1}$$

To simplify the notation, one can define a vector $\mathbf{f}^{(k)} = (f_1^{(k)}, \ldots, f_m^{(k)}) \in \mathbb{F}_p$, $\mathbf{F} = (F_1, \ldots, F_m) \in \mathbb{F}_p[X_1, \ldots, X_m]$ and the recurrence relation given by (1.1) becomes

$$\mathbf{f}^{(k+1)} = \mathbf{F}(\mathbf{f}^{(k)}), \ \forall k \ \text{ with } \ \mathbf{f}^{(0)} = \mathbf{X} = (X_1, \ldots, X_m). \tag{1.2}$$

In particular, denoting k applications of the recurrence relation by $F_i^{(k)}$, for any k, $n \geq 0$ and $i = 0, 1, \dots, m$

$$f_i^{(k+n)} = F_i^{(k)}(\mathbf{f}^{(n)}) = F_i^{(k+n)}(\mathbf{f}^{(0)})$$

and

$$\mathbf{f}^{(k+n)} = \mathbf{F}^{(k)}(\mathbf{f}^{(n)}) = \mathbf{F}^{(k+n)}(\mathbf{f}^{(0)}).$$

It is clear that the above sequence of vectors $f^{(k)}$ is eventually periodic with some period $T_v \leq p^m$ since it is generated over a finite field of p elements: that is,

$$\mathbf{f}^{(k+T_v)} = \mathbf{f}^{(k)},\ \forall k.$$

In the series of papers [18]-[25], multivariate polynomial systems $F_1, \dots, F_m$ of m polynomials in m variables over a finite field $\mathbb{F}_p$ are described in terms of the first iteration of (1.2), where the initial condition vector $\mathbf{f}^{(0)}$ is chosen as $\mathbf{X} = (X_1, \dots, X_m)$ and after the first iteration, entries of the vector $\mathbf{f}^{(1)} = \mathbf{F}(\mathbf{X}) = (F_1(\mathbf{X}), \dots, F_m(\mathbf{X}))$ are found as

$$F_1(\mathbf{X}) = X_1 G_1(X_2, \dots, X_m) + H_1(X_2, \dots, X_m),$$

$$F_2(\mathbf{X}) = X_2 G_2(X_3, \dots, X_m) + H_2(X_3, \dots, X_m),$$

$$\vdots$$

$$F_{m-1}(\mathbf{X}) = X_{m-1} G_{m-1}(X_m) + H_{m-1}(X_m),$$

$$F_m(\mathbf{X}) = g_m X_m + h_m, \tag{1.3}$$

with

$$G_i\,, H_i \in \mathbb{F}_p[X_{i+1}, \dots, X_m],\ i = 1, \dots, m-1$$

and

$$g_m,\ h_m \in \mathbb{F}_p,\ g_m \neq 0.$$

Table 1.1: Polynomial choices

Choice 1 in [21] $G_i(X_{i+1}, \ldots, X_m) = X_{i+1}$ and $H_i(X_{i+1}, \ldots, X_m) = h_i$ for $i = 1, \ldots, m-1.$
Choice 2 in [18]: $G_i(X_{i+1}, \ldots, X_m) = X_{i+1}^2 - a_i$ for some quadratic non-residues a_i and $H_i(X_{i+1}, \ldots, X_m) = h_i$ for $i = 1, \ldots, m-1.$
Choice 3 in [19]: $G_i(X_{i+1}, \ldots, X_m) = g_i, \;\; g_i, g_m \notin \{0, 1\}$ and Our choices for $H_i(X_{i+1}, \ldots,\; X_m)$ are: (a) $H_i = X_{i+1}$ (b) $H_i = X_{i+1}^2$ (c) $H_i = X_{i+1} \ldots X_m$ for $i = 1, \ldots, m-1.$
Choice 4 in [20]: $G_i(X_{i+1}, \ldots, X_m) = 1,$ and $H_i(X_{i+1}, \ldots,\; X_m) = X_{i+1}^{p-1} \ldots X_m^{p-1},$ for $i = 1, \ldots, m-1$ $g_m = 1, \; h_m \neq 0.$

The following iteration proceeds by substituting the obtained vector $\mathbf{f}^{(1)} = \mathbf{F}(\mathbf{X})$ instead of $\mathbf{X}$ in (1.3); so, $\mathbf{f}^{(2)} = \mathbf{F}(\mathbf{f}^{(1)})$. The structure of the polynomial description given by (1.3) is called the "triangular form", because it defines the first polynomial F_1 as a function of all the elements of $\mathbf{X}$, whereas the last polynomial F_m depends on a single element X_m of $\mathbf{X}$.

In order to obtain very fast pseudo-random generators, some choices for the polynomials are proposed by the same authors and they also propose polynomials in order to produce maximum-period vector sequences. For the system above, a first degree G_i polynomial with a constant H_i (Choice 1 in [21]), a second degree G_i polynomial with a constant H_i (Choice 2 in [18]), a constant G_i polynomial (Choice 3 in [19]) and $G_i = 1$ with a polynomial $deg(H_i) = (m - i)(p - 1)$ (Choice 4 in [20]). These polynomial choices for $i = 1, \ldots, m - 1$ are summarized in Table 1.1.

1.3 Testing Randomness

Unpredictability or randomness of a sequence is measured by its entropy [30]. In [7] as an answer to the question of "How are we to decide whether a sequence is sufficiently random?", the first collection of empirical randomness tests are given; i.e., Equidistribution test (Frequency test), Serial test, Gap test, Poker test (Partion test), Coupon collector's test, Permutation test, Run test, Maximum-of-t test, Collision test and Serial correlation test. CRYPT-X [4], DIEHARD Test Suite [12], NIST Test Suite [29] and TESTU01 [10] are the other test suits for measuring the randomness of sequences.

In this work, we analyze the randomness of sequences with respect to two basic approaches: distributions of their periods and linear complexities. Letting $\mathrm{s} = (s_1, \ldots, s_n)$ be a finite sequence of period T over $\mathbb{F}_p$, where $1 \leq n \leq 2T$, we measure its randomness according to the following properties:

- *Property* 1: The period T of the sequence should be sufficiently large; i.e., close to the maximum possible period ([16]).
- *Property* 2: Its linear complexity should be close to its period T and its linear complexity profile graph should be close to the $n/2$-line in its first two periods ([28]).

1.4 Aim and Skeleton of the Thesis

In this thesis, we analyze the randomness of the pseudo-random sequences produced by the multivariate polynomial iterations of Ostafe and Shparlinski [18]-[25]. For this purpose we investigate the distributions of the periods and linear complexities of the generated sequences by 6 different versions (so called Choices 1, 2, 3a, 3b, 3c and 4 given in Table 1.1) of the multivariate polynomial iterations method.

In Chapter 2, we perform an exhaustive search in order to obtain the period distribution of the vector sequences generated by the first five polynomial choices (Choice 1, 2, 3a, 3b and 3c) for prime field sizes p up to 13 and the vector sizes m up to 5. In Chapter 3, we exhaustively calculate the linear complexities of the sequences handled in the first chapter. In Chapter 4, the linear complexities of the sequences obtained at the highest periods of the corresponding choices are analyzed. In Chapter 5, we examine the linear complexity of the sequences produced by Choice 1, 2 and 3a by fixing the length of the sequence and using random vector size m together with random initial values. Then, we conclude the results of our study in the last chapter.

CHAPTER 2

EXHAUSTIVE PERIOD ANALYSIS

2.1 Introduction

In this chapter, we have exhaustively analyzed the vector periods of the sequences generated by the multivariate polynomial system (1.3) proposed by Ostafe and Shparlinski [18]-[25] with five polynomial choices: Choice 1, 2, 3a, 3b and 3c given in Table 1.1. Choice 4, generates the maximum-period sequences; i.e, the vector period T_v of the generated sequence is equal to p^m under all possible initial conditions; therefore, there is no need to include Choice 4 in the period analysis. In Section 2.2, we describe the parameter sets for these five polynomial choices. In Section 2.3, we consider ten cases in terms of the field size p and the number of polynomials m, and obtain the corresponding period distributions for finite field sizes p up to 13 and the number of polynomials m up to 5. In Section 2.4, we mention some observations about the factors of periods.

2.2 Set of Parameters

Our aim is to investigate the period distribution of the sequences generated by (1.3) for Choices 1, 2, 3a, 3b and 3c demonstrated in Table 1.1. Before performing an exhaustive search over all possible sequences, one needs to know the size of the parameter sets. Table 2.1 shows the size of the parameter sets and Table 2.2 depicts the number of possible vector sequences corresponding to each of the five choices, where Q_p and $\bar{Q}_p$ denote the number of quadratic residues and non-residues respectively, in mod p.

Table 2.1: Size of parameter sets

Choice	1	2	3a	3b	3c
X_i	p^m	p^m	p^m	p^m	p^m
G_i	-	-	$(p-2)^{m-1}$	$(p-2)^{m-1}$	$(p-2)^{m-1}$
H_i	p^{m-1}	p^{m-1}	-	-	-
a_i	-	$\bar{Q}_p^{m-1}$	-	-	-
g_m	$p-1$	$p-1$	$p-2$	$p-2$	$p-2$
h_m	p	p	p	p	p

Table 2.2: Number of possible vector sequences

Choice	Number of Possible Vector Sequences
1	$p^{2m}(p-1)$
2	$p^{2m}\bar{Q}_p^{m-1}(p-1)$
3a	$p^{m+1}(p-2)^m$
3b	$p^{m+1}(p-2)^m$
3c	$p^{m+1}(p-2)^m$

We examine 10 cases, corresponding to relatively small values of the field size p, and the vector size m ($p = 3$ with $2 \leq m \leq 5$; $p = 5, 7$ with $m = 2, 3$; and $p = 11, 13$ with $m = 2$), since for higher values of p and m, the behavior of the algorithms are predictable whereas the size of the exhaustive search space increases exponentially as shown in 2.2 (see Appendix A for search durations over a computer with Intel(R) Xeon(R) CPU 3.70 GHz). We choose the field size $p > 2$, since Choice 2, 3a, 3b and 3c do not work over $\mathbb{F}_2$ and Choice 1 can only generate very short sequences with $T_v \leq 2$ (see Appendix B for the proof). Total numbers of possible vector sequences for these 10 cases are shown in Table 2.3 as computed using Table 2.2, where the number of quadratic non-residues $\bar{Q}_3, \bar{Q}_5, \bar{Q}_7, \bar{Q}_{11}$ and $\bar{Q}_{13}$ are respectively equal to 1, 2, 3, 5 and 6.

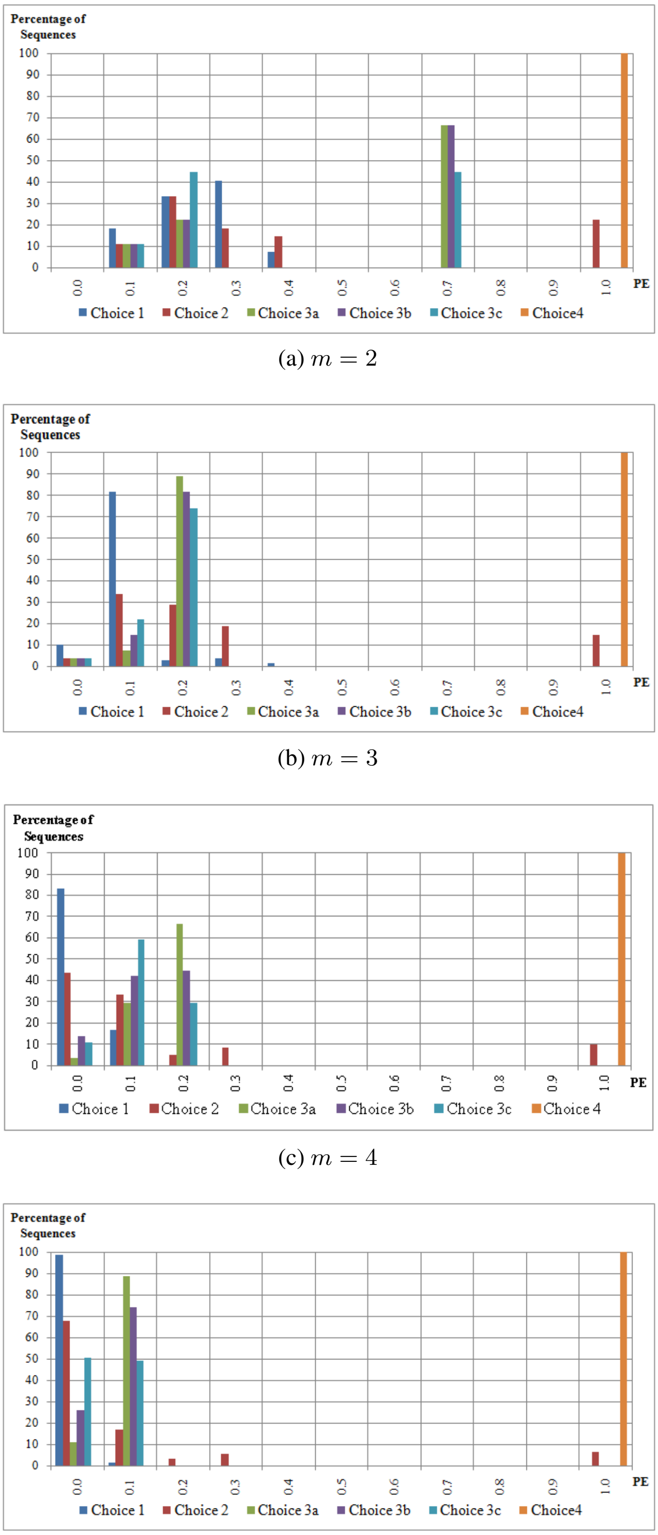
Table 2.3: Number of possible vector sequences for the ten cases

Case	p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
1	3	2	162	162	27	27	27
2	3	3	1458	1458	81	81	81
3	3	4	13122	13122	243	243	243
4	3	5	119466	118098	729	729	729
5	5	2	2500	5000	1125	1125	1125
6	5	3	62500	250000	16875	16875	16875
7	7	2	14406	43218	8575	8575	8575
8	7	3	705894	6353046	300125	300125	300125
9	11	2	146410	732050	107811	107811	107811
10	13	2	342732	2056392	265837	265837	265837

2.3 Period Distributions

We generate all vector sequences for 10 cases by Ostafe and Shparlinski's method [18]-[25], using the first five polynomial choices shown in Table 1.1. The last one, Choice 4 does not require any exhaustive search for periods, because it always generates the maximum period sequences with $T_v = p^m$. We then compute the vector period T_v of each vector sequence, and find how close it is to the maximum period by computing its "Period Efficiency, PE $= T_v/p^m$". Corresponding period efficiency distributions are sketched in Figures 2.1-2.3 for the six polynomial choices over all possible values of $\mathbf{X} = (X_1, ..., X_m)$, $\mathbf{G} = (G_1, ..., G_{m-1})$, $\mathbf{H} = (H_1, ..., H_{m-1})$, a_i, g_m and h_m, having the set sizes given in Table 2.1. In all figures, we indicate the period distribution corresponding to Choice 1, 2, 3a, 3b, 3c and 4 (as a reference).

It can be observed from Figures 2.1-2.3 that the period efficiencies of Choice 1 sequences do not exceed 0.5 for all 10 cases. As seen in Figure 2.1, for $p = 3$, the distributions of PE's generally move towards the origin as m increases from 2 to 4, but for $m = 5$, some high-PE Choice 3 sequences appear around PE $= 0.85$. Figure 2.2 gives an idea about the change of PE distributions versus the field size p for $m = 2$ (as also compared to $m = 2$ case of the previous figure), where one clearly spots the general trend of decreasing T_v's that move towards the origin as p increases, again with the exception of some high-PE Choice 3 sequences that show up around PE $= 0.95$ for $p = 11$. Finally, Figure 2.3 indicates that when the number of polynomials, m, increases from 2 to 3, for $p = 5$ and 7, one cannot observe a net movement in the PE distributions towards the origin as sharp as that of the $p = 3$ case seen in Figure 2.1. However, one can also not say that an increase in m improves the PE distribution.



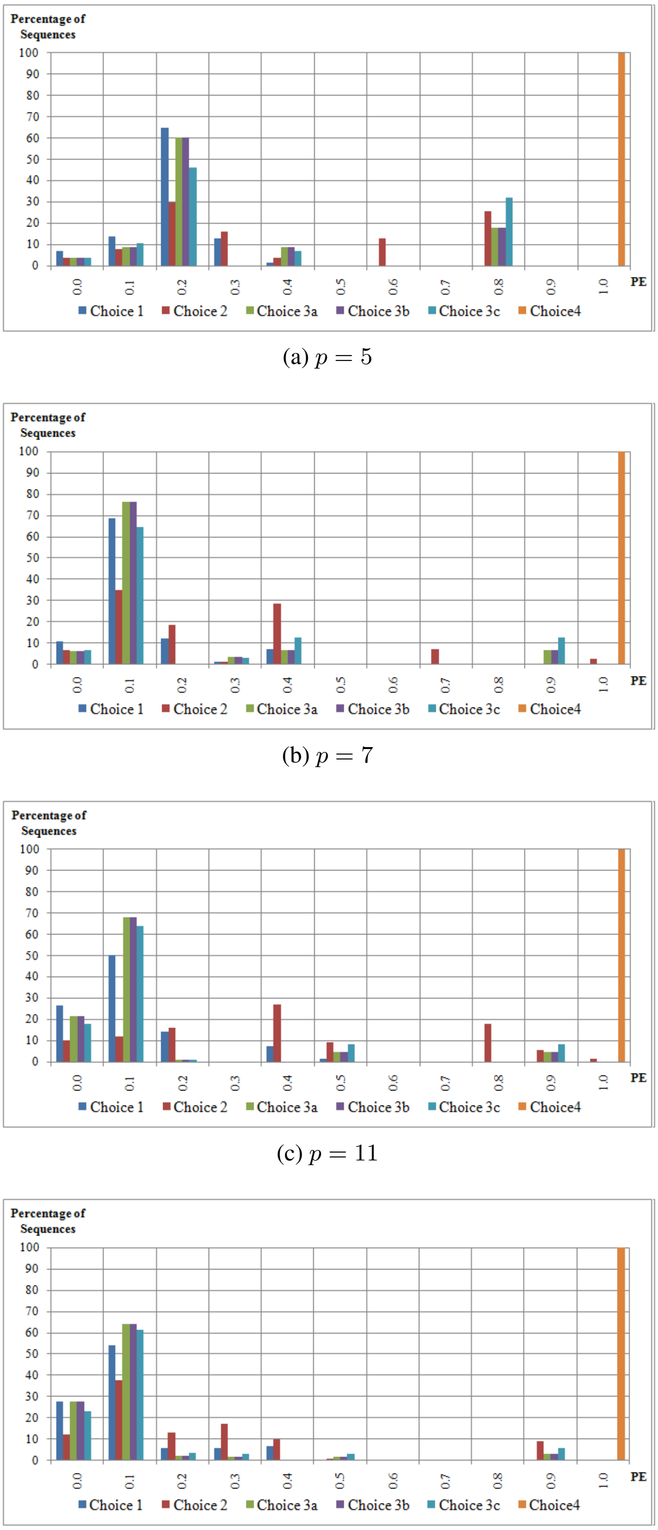
(a) $m = 2$

(b) $m = 3$

(c) $m = 4$

(d) $m = 5$

Figure 2.1: Distribution of the period efficiency, PE (the period T_v of the vector sequence divided by the maximum period p^m), for Choice 1, 2, 3a, 3b, 3c and 4 and the field size $p = 3$, where the number of polynomials are: (a) $m = 2$, (b) $m = 3$, (c) $m = 4$, (d) $m = 5$



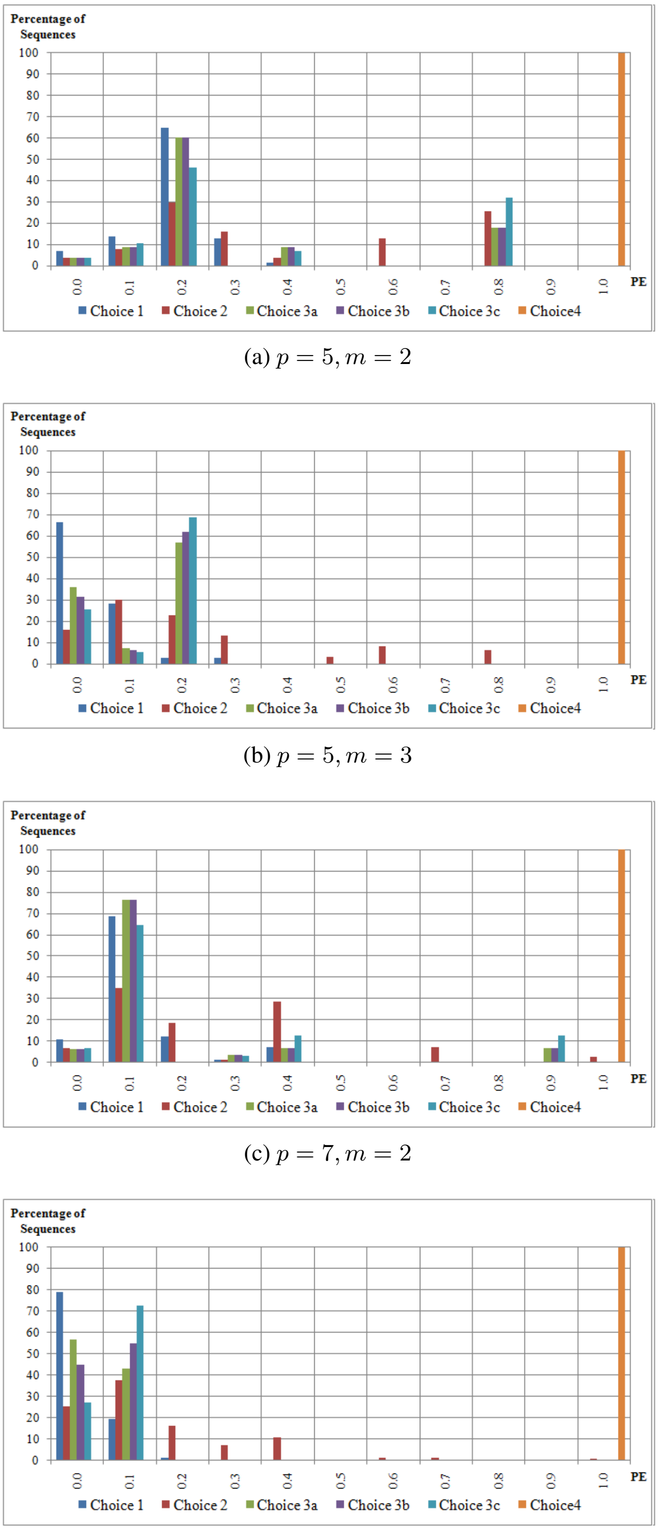
(a) $p = 5$

(b) $p = 7$

(c) $p = 11$

(d) $p = 13$

Figure 2.2: Distribution of the period efficiency, PE (the period T_v of the vector sequence divided by the maximum period p^m), for Choice 1, 2, 3a, 3b, 3c and 4 and the field sizes: (a) $p = 5$, (b) $p = 7$, (c) $p = 11$, (d) $p = 13$ and the number of polynomials $m = 2$



(a) $p = 5, m = 2$

(b) $p = 5, m = 3$

(c) $p = 7, m = 2$

(d) $p = 7, m = 3$

Figure 2.3: Distribution of the period efficiency, PE (the period T_v of the vector sequence divided by the maximum period p^m), for Choice 1, 2, 3a, 3b, 3c and 4; where the field size p and the number of polynomials m are given as (a) $p = 5, m = 2$, (b) $p = 5, m = 3$, (c) $p = 7, m = 2$, (d) $p = 7, m = 3$

Table 2.4: Weighted average of PE values for ten cases

p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
3	2	23.70	41.48	52.22	52.22	41.11
3	3	10.49	29.63	18.52	17.78	17.04
3	4	1.68	16.69	16.30	13.09	11.85
3	5	0.13	10.57	8.89	7.41	4.94
5	2	18.80	41.28	30.76	30.76	38.76
5	3	4.12	23.22	12.09	13.03	14.32
7	2	12.45	26.68	17.58	17.58	23.90
7	3	2.21	15.74	4.31	5.49	7.27
11	2	11.59	40.89	13.33	13.33	18.61
13	2	11.17	24.17	10.83	10.83	14.48

Table 2.5: Percentage of Choice 2 sequences with maximum-period p^m

p	3	3	3	3	5	5	7	7	11	13
m	2	3	4	5	2	3	2	3	2	2
Percentage	22	15	10	7	0	0	3	1	2	0

Table 2.4 shows the weighted average of PE values for each case by multiplying the PE values with corresponding percentages and summing over all PE values. As can be seen from Table 2.4, increasing the field size p and the vector size m decreases the weighted average of PE values almost for all choices with the single exception of Choice 2 for $p = 11$.

Figure 2.1-2.3 also show that Choice 1, 3a, 3b and 3c do not produce any maximum-period ($T_v = p^m$) vector sequence for the considered parameters ($p = 3$ with $2 \leq m \leq 5$; $p = 5, 7$ with $m = 2, 3$; and $p = 11, 13$ with $m = 2$). In addition to Choice 4, generation of sequences with maximum periods seems to be possible with Choice 2 as well; however, the corresponding percentages are very small. We tabulate the percentage of maximum-period sequences for Choice 2 in Table 2.5 that is observed to be less than %3 if $3 < p \leq 13$. As mentioned above, the last choice in Table 1.1, Choice 4, always produces sequences at maximum period, with PE $= 1$. It is of further interest to find whether the highest-period sequences of each choice are random enough; which is the subject of Chapter 4.

Table 2.6 shows that the period efficiency T_v/p^m of the generated vector sequences are less than or equal to 0.5 for all sequences generated by Choice 1, for more than 62% of Choice 2 sequences, for more than 82% of Choice 3a and 3b sequences (except for $p = 3, m = 2$ with 33%) and for more than 56% of Choice 3c sequences.

Table 2.6: Percentage of sequences with period efficiency ≤ 0.5

Case	p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
1	3	2	100	78	33	33	56
2	3	3	100	85	100	100	100
3	3	4	100	90	100	100	100
4	3	5	100	93	100	100	100
5	5	2	100	62	82	82	68
6	5	3	100	85	100	100	100
7	7	2	100	90	93	93	87
8	7	3	100	97	100	100	100
9	11	2	100	75	96	96	91
10	13	2	100	91	97	97	94

2.4 Factors of Periods

Although the generation of high-period sequences by the first five polynomial choices of Table 1.1 seem to be less probable than low-period sequences, one may also be interested in examining the whole set of possible periods that can be generated. We tabulate all possible periods generated in the 10 considered cases in Appendix C. Table 2.7 parameterizes the highest periods found by exhaustive search for the examined 10 cases, using the further details given in Tables C.3-C.6 of Appendix C.

Similar to Theorem 8 in [20], we have observed the following fact about the period of the vector sequence: Each period $T_v = t_1 ... t_m$ is the product of m integers $t_1 ... t_m$, which can be equal to p or to a factor of $p - 1$ for all five choices. More specifically, for $m = 2$, each period $T_v = t_1 t_2$ is the product of two integers t_1 and t_2, which can be equal to p or to a factor of $p - 1$. Similarly, each period for $m = 3$ is the product of three integers, for $m = 4$, it is the product of four integers, and for $m = 5$, it is the product of five integers that can be equal to p, or to a factor of $p - 1$. In Appendix C, one can also examine that Choice 2 generates the largest set of period values, which almost always contains Choice 1, 3a, 3b and 3c sets and some extra values. On the other hand, Choice 3a, 3b and 3c have the smallest set of periods in all choices.

Table 2.7: The highest vector periods of the corresponding polynomial choices

m	p	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
2	$3 \leq p \leq 13$	$p(p-1)/2$ $[(p-1)^2\,]$ only for $p=3$	$p(p-1)$ or p^2	$p(p-1)$	$p(p-1)$	$p(p-1)$
3	$3 \leq p \leq 7$	$p(p-1)^2$ or $p(p-1)^2/2$ or $p(p-1)^2/3$	$p^2(p-1)$ or p^3	$p(p-1)$	$p(p-1)$	$p(p-1)$
4	3	$p(p-1)^2$	p^4	$p^2(p-1)$	$p^2(p-1)$	$p^2(p-1)$
5	3	$p^2(p-1)^2$	p^5	$p^2(p-1)$	$p^2(p-1)$	$p^2(p-1)$

2.5 Conclusion

Our exhaustive search on the distribution of the vector period T_v generated by (1.3), for 10 cases ($p = 3$ with $2 \leq m \leq 5$; $p = 5, 7$ with $m = 2, 3$; and $p = 11, 13$ with $m = 2$), shows that there is no maximum-period sequence (maximum possible period of the scalar sequence is $T_{max} = mp^m$) for Choice 1 [21] and Choice 3a, 3b, 3c [19] given in Table 1.1. Choice 2 [18] is more promising since it has smaller percentage of small-period sequences than other choices; and maximum-period sequences with period mp^m do exist, although their existence probability is less than 3% if $p > 3$, as can be observed in Table 2.5. All sequences generated by Choice 1, have very low period efficiencies, PE $= T/mp^m$, less than or equal to 0.5 (See Table 2.6). Similarly, the percentage of low period sequences with period efficiency less than 0.5 is more than 62% for Choice 2, more than 82% for Choice 3a and 3b (except for $p = 3, m = 2$ with 33%) and more than 56% for Choice 3c sequences.

CHAPTER 3

EXHAUSTIVE LINEAR COMPLEXITY ANALYSIS

3.1 Introduction

After examining the period distributions of the sequences produced by Choice 1, 2, 3a, 3b and 3c exhaustively; we perform another exhaustive search in this chapter, again over all possible initial conditions within the sets whose sizes are as given in Table 2.3, in order to find the linear complexity distributions of the sequences produced by (1.3) for ten specific cases of the field size p and the vector size m. Section 3.2 is devoted to some essential definitions about the linear complexity. In Section 3.3, we describe our analysis method to measure the linear complexity and introduce a parameter that we call "the linear complexity efficiency, LCE", taking real values in the interval [0, 1]. In Section 3.4, we sketch the linear complexity distributions. In Section 3.5, we find the minimum, average and maximum LCE's of the sequences versus their vector periods, T_v.

Then, we consider some subsets of practical significance chosen from the overall space. First, for the sequences whose periods are at least half of the maximum possible period, we present the minimum, average and maximum linear complexity values in Section 3.6. Finally, we concentrate on sequences with high linear complexities (LCE ≥ 0.95) and Section 3.7 presents the computed percentages of such sequences together with their vector periods.

Discussion of Choice 4 that only produces sequences with maximum periods is left to Chapter 4, which is devoted to the LCE investigation of largest-period sequences produced by all choices given in Table 1.1.

3.2 Linear Complexity

This section is intended to give some fundamental definitions [17] on the concept of linear complexity.

Definition 3.1. *The sequence $s_1, s_2, ...$ over $\mathbb{F}_q$ satisfies a linear recurrence relation over $\mathbb{F}_q$ of order k if there exist $a_0, a_1, ..., a_{k-1} \in \mathbb{F}_q$ such that*

$$s_{i+k} = \sum_{h=0}^{k-1} a_h s_{i+h} \ \ for \ \ i = 1, 2, ...$$

where k is a positive integer.

The linear recurrence relation and the initial values $s_1, ..., s_k$ uniquely determine the sequence $s_1, s_2, ...$ given in Definition 3.1. Definition 3.2 and 3.3 clarify the linear complexity and the linear complexity profile, respectively.

Definition 3.2. *Let S be either a finite or an infinite duration sequence over $\mathbb{F}_q$ containing at least n terms, and n be a positive integer. Then the n-th linear complexity $L_n(S)$ of S is the smallest k, for which a linear recurrence relation over $\mathbb{F}_q$ of order k can generate the first n terms of S. If S is ultimately periodic, then its linear complexity $L(S)$ is defined by*

$$L(S) = \sup_{n \geq 1} L_n(S).$$

Definition 3.3. *Let $L_n(S)$ denote the n-th linear complexity of an infinite sequence S over $\mathbb{F}_q$. Then, the sequence $L_1(S), L_2(S), ...$ is called the linear complexity profile of S.*

By using the algorithm invented by Berlekamp and Massey ([1], [14]) one can recover the linear recurrence relation and the initial values from the first $2k$ terms of the sequence. The all zero sequence $(0, 0, ...)$ over $\mathbb{F}_q$ satisfies a linear recurrence relation over $\mathbb{F}_q$ of order 0, by convention.

3.3 Measuring Randomness in Terms of the Linear Complexity

Our aim is to investigate the randomness of the scalar sequences generated by (1.3) for the five choices of the polynomials (Choice 1, 2, 3a, 3b and 3c) given in Table 1.1, in terms of the linear complexity. We compute the linear complexity of a sequence, by using the Berlekamp-Massey algorithm ([1], [14]). In order to measure whether the linear complexities of the scalar sequences are close to their period T or not (period T of the scalar sequence is equal to the product of the number of polynomials m and the period T_v of the vector sequence), we define a criterion called linear complexity efficiency (LCE), L/T as the "ratio of the computed linear complexity L of the sequence to its length T". In order to satisfy the second randomness property stated in Section 1.3, a sequence should have an LCE value close to 1.

Table 3.1: Weighted average of LCE values for ten cases

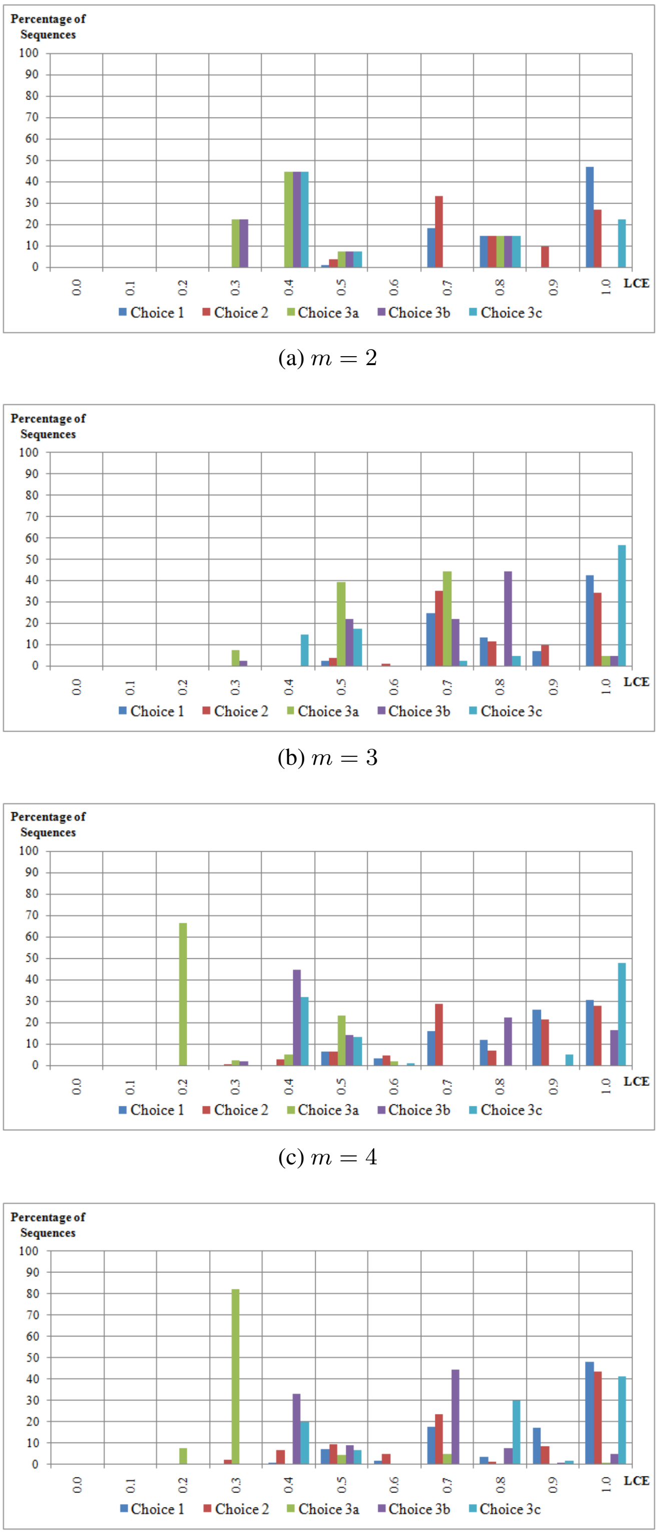
p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
3	2	72.35	73.09	40.00	40.00	55.56
3	3	77.80	79.78	58.02	67.90	77.04
3	4	79.59	79.71	28.56	59.51	72.10
3	5	83.65	79.30	32.37	60.27	78.05
5	2	72.94	52.58	48.39	48.39	49.24
5	3	77.52	63.57	39.34	42.23	46.10
7	2	70.50	57.14	41.45	41.45	45.49
7	3	76.72	46.06	35.85	35.77	32.17
11	2	63.17	32.33	30.67	30.67	34.92
13	2	59.23	44.33	32.70	32.70	33.48

3.4 Linear Complexity Distributions

We have exhaustively computed the linear complexity efficiencies of sequences generated by (1.3) with the first five polynomial choices given in Table 1.1 for 10 cases of the field size p, and the vector size m (i.e., $p = 3$ with $2 \leq m \leq 5$; $p = 5, 7$ with $m = 2, 3$; and $p = 11, 13$ with $m = 2$) over all possible values $\mathbf{X} = (X_1, ..., X_m)$, $\mathbf{G} = (G_1, ..., G_{m-1})$, $\mathbf{H} = (H_1, ..., H_{m-1})$, a_i, g_m and h_m given in Table 2.1. As mentioned in Appendix C, the sequences with vector period $T_v = 1$ are discarded, since they reflect the randomness of the initial vector $\mathbf{X} = (X_1, ..., X_m)$ rather than that of the multivariate polynomial iterations method given by (1.3). Corresponding linear complexity efficiency distributions are sketched in Figures 3.1-3.3 for the six polynomial choices over all possible values of $\mathbf{X} = (X_1, ..., X_m)$, $\mathbf{G} = (G_1, ..., G_{m-1})$, $\mathbf{H} = (H_1, ..., H_{m-1})$, a_i, g_m and h_m, having the set sizes given in Table 2.1 for all choices.

Table 3.1 shows the weighted average of LCE values for each case obtained by multiplying the LCE values with corresponding percentages and summing over all LCE values.

As can be seen from Table 3.1, increasing the vector size m increases the weighted average LCE of the sequences generated by Choice 1. It seems that Choice 2 is worse than Choice 1. Another observation is that weighted average LCE is decreasing, while m is increasing for Choice 3a. Choice 3a is the worst of all choices in terms of the linear complexity. On the other hand, distributions of Choice 3b and 3c are quite similar.



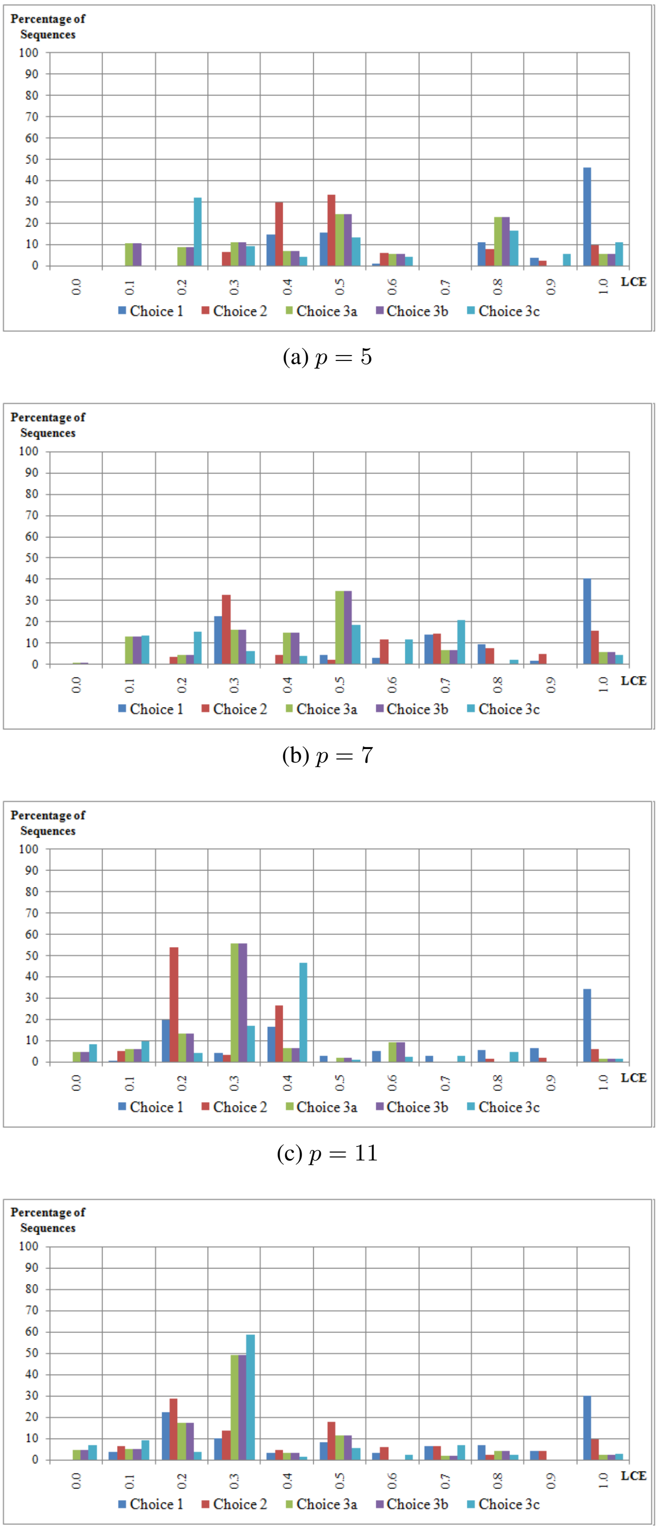
(a) $m = 2$

(b) $m = 3$

(c) $m = 4$

(d) $m = 5$

Figure 3.1: Distribution of the linear complexity efficiency, LCE (ratio of the computed linear complexity L of the sequence to its length T), for Choice 1, 2, 3a, 3b and 3c (respectively dark blue, red, green, purple, light blue and orange) and the field size $p = 3$, where the number of polynomials are: (a) $m = 2$, (b) $m = 3$, (c) $m = 4$, (d) $m = 5$



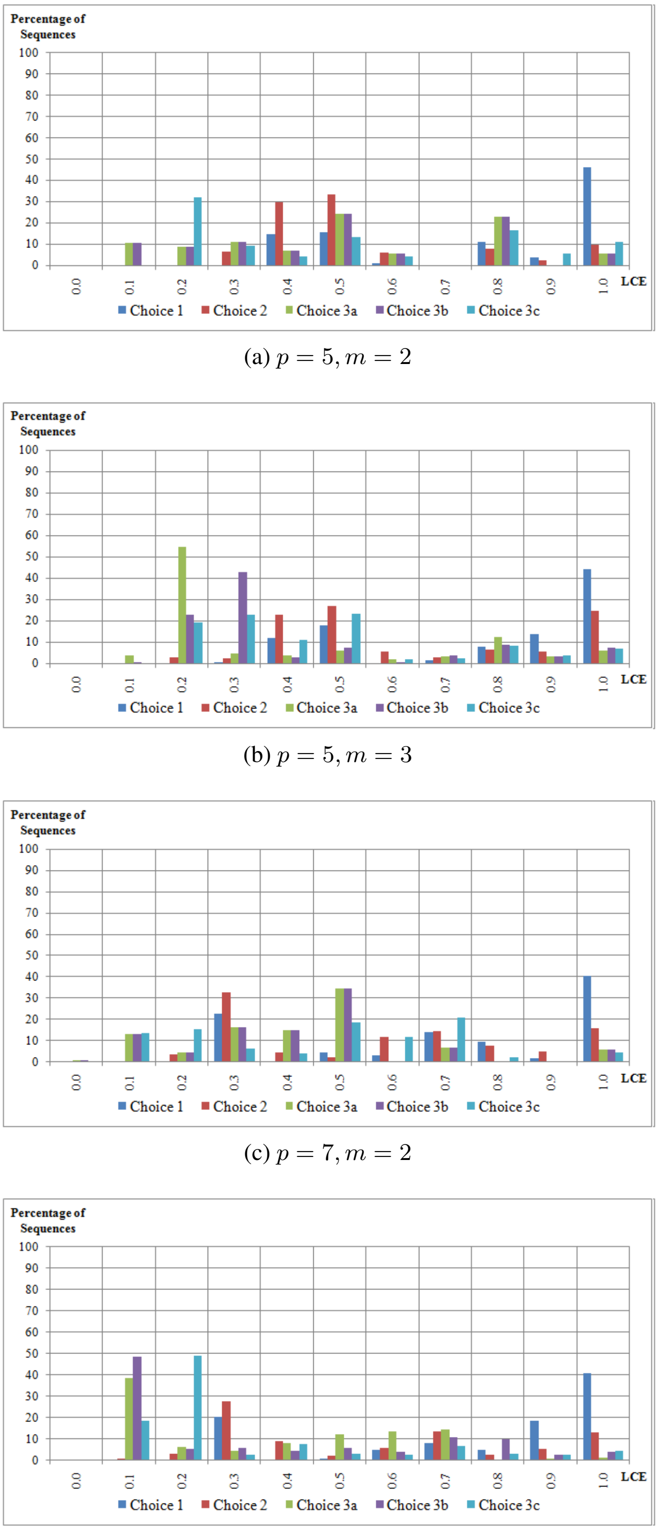
(a) $p = 5$

(b) $p = 7$

(c) $p = 11$

(d) $p = 13$

Figure 3.2: Distribution of the linear complexity efficiency, LCE (ratio of the computed linear complexity L of the sequence to its length T), for Choice 1, 2, 3a, 3b and 3c and the field sizes: (a) $p = 5$, (b) $p = 7$, (c) $p = 11$, (d) $p = 13$ and the number of polynomials $m = 2$



(a) $p = 5, m = 2$

(b) $p = 5, m = 3$

(c) $p = 7, m = 2$

(d) $p = 7, m = 3$

Figure 3.3: Distribution of the linear complexity efficiency, LCE (ratio of the computed linear complexity L of the sequence to its length T), for Choice 1, 2, 3a, 3b and 3c; where the field size p and the number of polynomials m are given as (a) $p = 5, m = 2$, (b) $p = 5, m = 3$, (c) $p = 7, m = 2$, (d) $p = 7, m = 3$

3.5 Minimum, Average and Maximum LCE versus T_v of the Sequences Generated by the Five Polynomial Choices

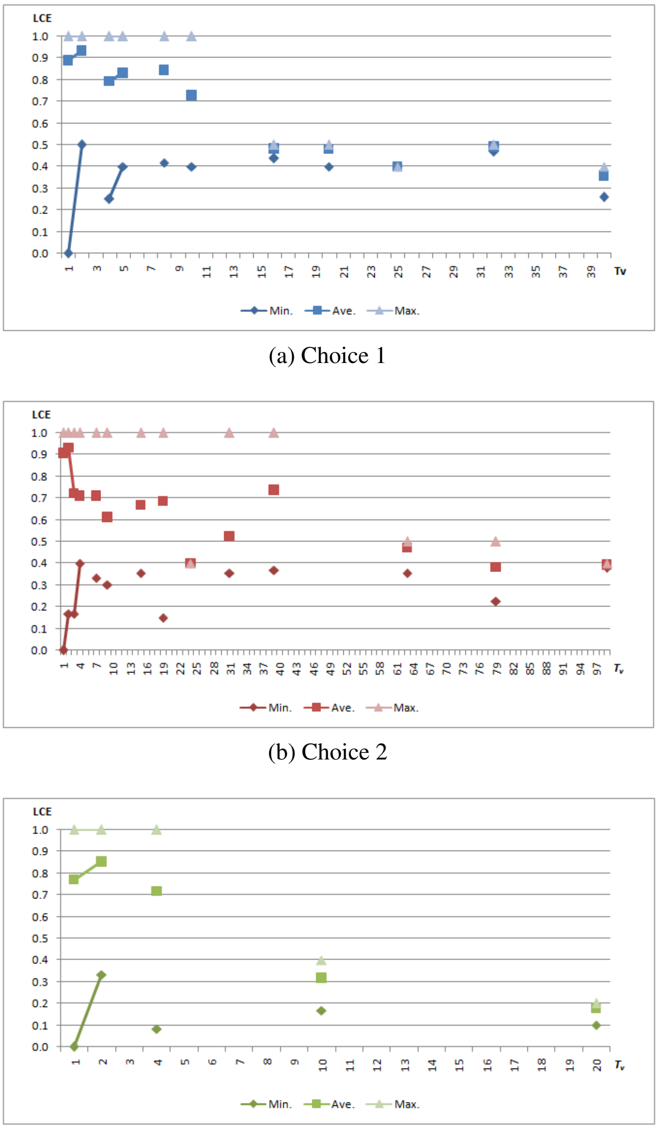
In order to examine the randomness of multivariate polynomial iterations method given by (1.3), we present the LCE values of the first five polynomial choices given in Table 1.1 versus the period T_v of the generated vector sequences for $T_v > 1$. Our exhaustive analysis is performed for 50 different cases (resulting from the product of 5 polynomial choices for each one of the 10 (p, m) pairs), and we present the most representative cases that include the largest variety of produced vector sequence periods; namely the (p, m) pairs of (5, 3), (7, 3), (11, 2), and (13, 2) in Figures 3.4-3.7 respectively. In each figure, we plot the minimum, average and maximum LCE values versus the vector sequence period T_v, corresponding to the polynomial Choices 1, 2 and 3a. We don't include Choice 3b and 3c, since they yield very similar curves to those of Choice 3a.

General characteristics of all these four figures are quite similar: (i) an increase in T_v results in serious loss of randomness for the three polynomial choices 1, 2 and 3a, (ii) Choice 2 produces the largest set of periods, followed by Choice 1 and Choice 3 (in accordance with the results of Chapter 2, as detailed by Tables C.3-C.6, (iii) an increase in the field size p also seems to yield some loss of randomness on the average. These observations are not encouraging for the practical use of Ostafe and Shparlinski's multivariate polynomial iterations method as a pseudo-random noise generator. Tables related to the details in Figures 3.4-3.7 are presented in D.1-D.12.

3.6 Linear Complexity Efficiency of Sequences with Period Efficiency, PE ≥ 0.5

The subset of sequences, with periods at least as large as one half of the maximum possible period, is of special interest. So, we inspect the linear complexity efficiencies of sequences having period efficiencies greater than or equal to 0.5. Table 3.2 shows the average LCE values of the sequences with PE ≥ 0.5 (also see Tables D.13 and D.14 for the minimum and maximum values of the LCE).

One can observe from Table 3.2 that average LCE values achieved by the sequences with PE ≥ 0.5 are decreasing while the field size p and the vector size m is increasing for all polynomial choices. As compared to the average LCE's obtained in the complete sets given in the previous section, one can say that high linear complexity efficiencies can not be seen for the sequences having period efficiencies ≥ 0.5.

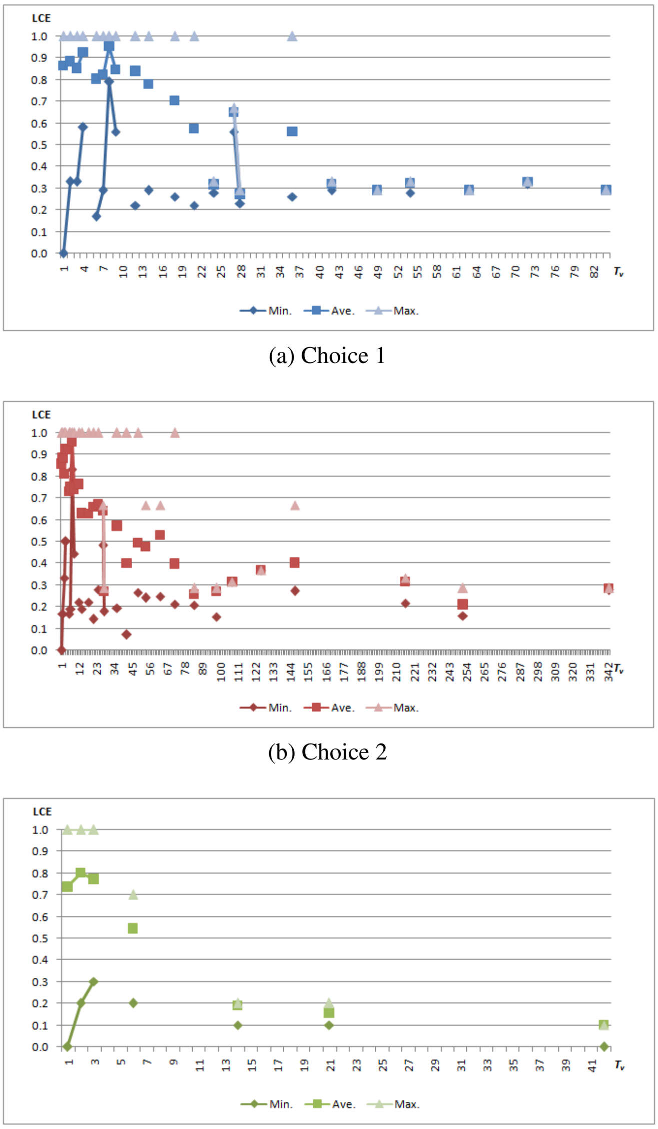


(a) Choice 1

(b) Choice 2

(c) Choice 3a

Figure 3.4: LCE values versus vector period T_v of sequences for $p = 5, m = 3$ generated by (a) Choice 1, (b) Choice 2, (c) Choice 3a polynomials

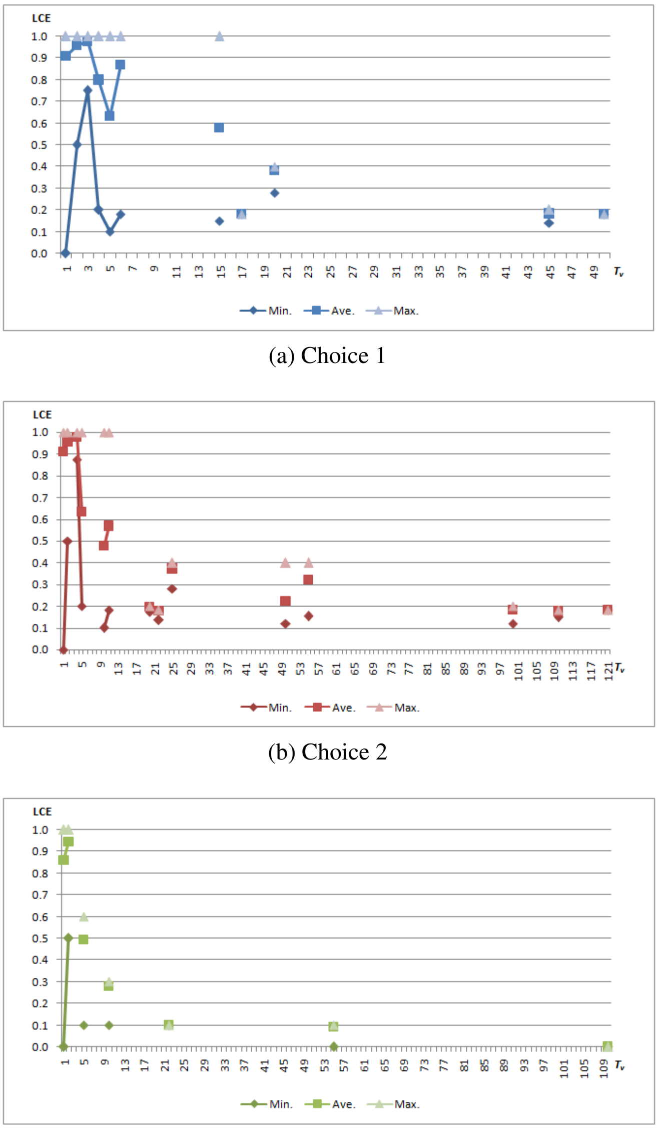


(a) Choice 1

(b) Choice 2

(c) Choice 3a

Figure 3.5: LCE values versus vector period T_v of sequences for $p = 7, m = 3$ generated by (a) Choice 1, (b) Choice 2, (c) Choice 3a polynomials

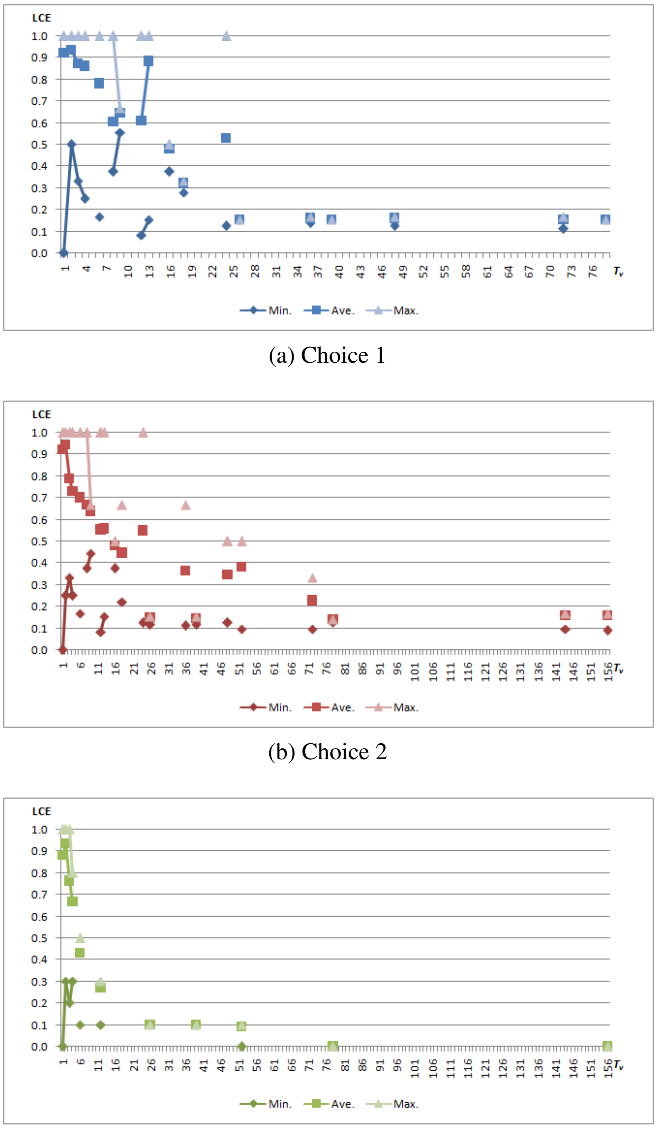


(a) Choice 1

(b) Choice 2

(c) Choice 3a

Figure 3.6: LCE values versus vector period T_v of sequences for $p = 11, m = 2$ generated by (a) Choice 1, (b) Choice 2, (c) Choice 3a polynomials



(a) Choice 1

(b) Choice 2

(c) Choice 3a

Figure 3.7: LCE values versus vector period T_v of sequences for $p = 13, m = 2$ generated by (a) Choice 1, (b) Choice 2, (c) Choice 3a polynomials

Table 3.2: Average linear complexity efficiency of the sequences with period efficiency ≥ 0.5, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m and h_m

Case	p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
1	3	2	-	0.67	0.39	0.39	0.42
2	3	3	-	0.67	-	-	-
3	3	4	-	0.67	-	-	-
4	3	5	-	0.67	-	-	-
5	5	2	-	0.43	0.14	0.13	0.18
6	5	3	-	0.40	-	-	-
7	7	2	-	0.30	0.09	0.07	0.09
8	7	3	-	0.27	-	-	-
9	11	2	0.18	0.22	0.05	0.04	0.04
10	13	2	0.15	0.16	0	0.02	0.02

3.7 Sequences with High Linear Complexity Efficiency, LCE ≥ 0.95

3.7.1 Percentages

Since one of the desired properties of randomness for a sequence is to have a high linear complexity as mentioned in Section 1.3, we calculate the percentage of sequences with high LCE, found over all possible values of parameters ($\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m, and h_m presented in Table 2.1) and illustrate them in Table 3.3.

One can observe from Table 3.3 that Choice 1 seems to generate sequences with LCE's ≥ 0.95 more efficiently than Choices 2, 3a and 3b. However, Choice 3c has higher percentages of high-LCE sequences for $p = 3$ and $m = 3, 4$. Within the specified set of p and m values, the percentage of the high-LCE Choice 1 sequences varies in the interval between 30% and 48%; and this percentage decreases with increasing p. Besides, at most 44% of the sequences generated by Choice 2 have LCE ≥ 0.95, corresponding to the case of $p = 3$, $m = 5$. This choice seems more inefficient for higher values of p, at which smaller percentages of high-LCE sequences are produced. Additionally, Choice 3a hardly generates sequences with LCE ≥ 0.95, and corresponding percentages do not exceed 6%. Like Choice 3a, Choice 3b also rarely generates sequences with LCE ≥ 0.95. The case of $p = 3$, $m = 4$ has 16% high-LCE sequences, which is the highest percentage for this choice. On the other hand, Choice 3c produces more sequences with high LCE (≥ 0.95) than Choices 3a and 3b. Hence, using H_i polynomials with higher degrees than one (as in Choice 3a) or two (as in Choice 3b) seems to increase the linear complexity efficiency. Especially for $p = 3$, Choice 3c seems more efficient than Choice 1 for $p = 3$ and $m = 3, 4$ as well.

Table 3.3: Percentage of sequences with linear complexity efficiency LCE ≥ 0.95, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m and h_m

Case	p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
1	3	2	47	27	0	0	22
2	3	3	42	34	5	5	57
3	3	4	30	28	0	16	48
4	3	5	48	44	1	5	41
5	5	2	46	10	6	6	11
6	5	3	44	25	6	7	7
7	7	2	40	16	6	6	4
8	7	3	41	13	1	4	4
9	11	2	34	6	1	1	1
10	13	2	30	10	2	2	3

3.7.2 Corresponding Vector Periods

The other interesting question is “Which vector periods are achieved by the sequences having high LCE values?”. The answer of the question is presented in Table 3.4. It shows the vector periods of sequences with linear complexity efficiencies greater than or equal to 0.95, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m, and h_m.

It can be observed from Table 3.4, that all five polynomial choices produce less sequences with LCE ≥ 0.95 at high vector periods than at low vector periods.

3.8 Conclusion

The exhaustive search of LCE values over all possible initial conditions for $p = 3$ with $m = 3, 4, 5$, $p = 5, 7$ with $m = 2, 3$, and $p = 11, 13$ with $m = 2$ shows that none of Ostafe’s multivariate polynomial iterations generate sequences having good randomness properties. For all polynomial choices (1, 2, 3, 3b and 3c), Table 3.2 shows that LCE values achieved by the sequences with PE ≥ 0.5 are decreasing, while the field size p and the vector size m are increasing. We observe from Table 3.3 that Choice 1 seems to generate sequences with LCE’s ≥ 0.95 more efficiently than Choices 2, 3a and 3b. Choice 3a and 3b hardly generate sequences with LCE ≥ 0.95, and corresponding percentages do not exceed 6% and 16%, respectively. We also observe from Table 3.4 that all five polynomial choices produce less sequences with LCE ≥ 0.95 at high vector periods than at low vector periods and an increase in T_v results in serious loss of randomness.

Table 3.4: Vector periods of sequences with linear complexity efficiency LCE ≥ 0.95, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m and h_m

Case	p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
1	3	2	1, 2, 3, 4	1, 2, 3, 4	-	-	2
2	3	3	1, 2, 3, 4, 6	1, 2, 3, 4, 8, 9	1, 2	1, 2	1, 2, 6
3	3	4	1, 2, 3, 4, 6, 8, 12	1, 2, 3, 4, 8, 9, 27	-	1, 2, 6	1, 2, 6
4	3	5	1, 2, 3, 4, 6, 8, 9, 12	1, 2, 3, 4, 8, 9, 16, 27, 81	1, 2	6	2, 6
5	5	2	1, 2, 4, 5, 8	1, 2, 4, 5	1, 2	1, 2	1, 2, 4
6	5	3	1, 2, 4, 5, 8, 10	1, 2, 4, 5, 8, 10, 16, 20, 32, 40	1, 2, 4	1, 2, 4	1, 2, 4
7	7	2	1, 2, 3, 4, 6, 7, 12	1, 2, 3, 4, 6, 7, 12	1, 2, 3	1, 2, 3	1, 2, 3
8	7	3	1, 2, 3, 4, 6, 7, 8, 9, 12, 14, 18, 21, 36	1, 2, 3, 4, 6, 7, 8, 9, 12, 14, 18, 21, 24, 36, 42, 49, 72	1, 2, 3	1, 2, 3, 6	1, 2, 3, 6
9	11	2	1, 2, 4, 5, 10, 11, 20	1, 2, 4, 5, 10, 11	1, 2	1, 2	1, 2
10	13	2	1, 2, 3, 4, 6, 8, 12, 13, 24	1, 2, 3, 4, 6, 8, 12, 13, 24	1, 2, 3	1, 2, 3	1, 2, 3, 4

CHAPTER 4

LINEAR COMPLEXITY ANALYSIS OF THE LARGEST PERIOD SEQUENCES

4.1 Introduction

In this chapter we fix our attention to the largest period sequences that can be produced for a given polynomial choice, at a specific p and m. All computed LCE's, and their average, maximum, minimum values are found within these sets of the largest period sequences. Section 4.2 investigates the variation of the LCE values versus the field size p. In Section 4.3, we examine the variation of LCE values in a field with $p = 3$, versus the vector size m. The remaining sections are devoted to Choice 4, which always produces maximum-length sequences of vector period p^m. After reviewing the details of Choice 4 in Section 4.4, we compute the linear complexity of Choice 4 sequences using the Berlekamp-Massey algorithm and also present an example that demonstrates the poor randomness of the sequences produced by Choice 4.

4.2 LCE's Obtained at the Largest T_v of the Corresponding Choices versus the Field Size p

In order to observe the variation of the linear complexity efficiency of the largest period sequences versus the field size p, we consider p values up to 13 setting the vector size m equal to 2. For each of the polynomial choices 1, 2, 3a, 3b and 3c stated in Table 1.1, we exhaustively produce all possible sequences with the largest vector period T_v, and then use the Berlekamp-Massey algorithm to compute the linear complexities of all sequences in this set.

Figure 4.1 shows the variation of minimum, average and maximum linear complexity efficiencies attained within the set of the largest period sequences of the corresponding polynomial choices (1, 2, 3a, 3b and 3c) versus the field size $3 \leq p \leq 13$ with the number of polynomials, $m = 2$.

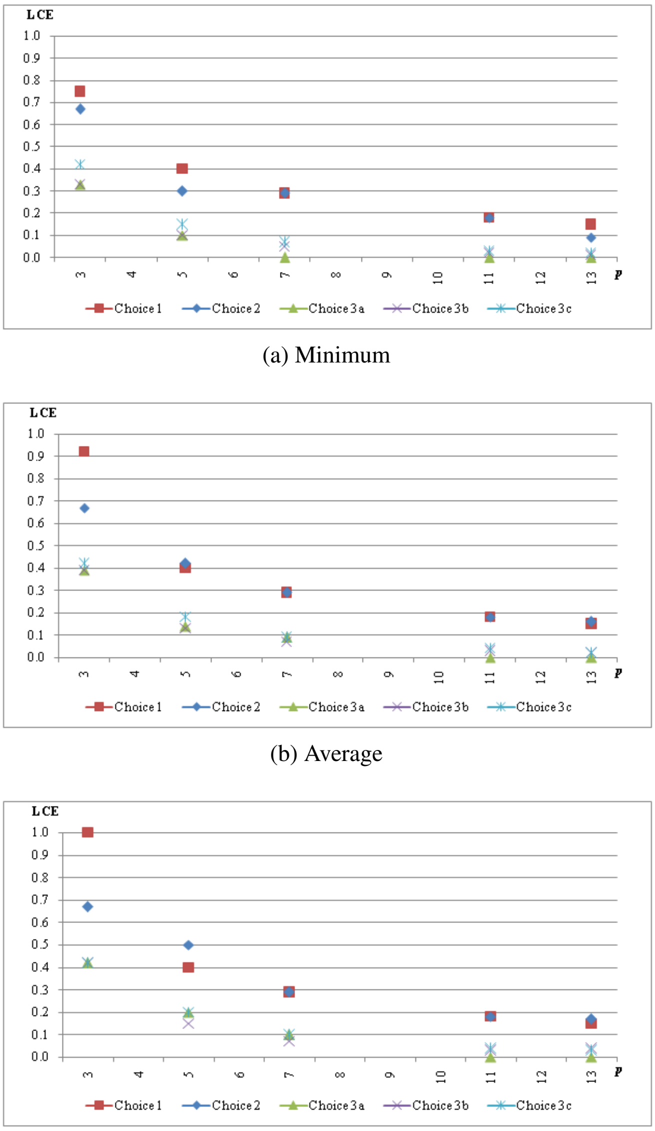
As can be seen from Figure 4.1, as the field size p increases from 3 to 13, all five polynomial choices generate sequences with lower and still lower LCE values. Although for the vector size $m = 2$, Choice 1 sequences seem slightly better than Choice 2 sequences, which are much better than Choice 3 sequences; yet none of the LCE values in these figures exceed 0.95, except that of Choice 1 for $p = 3$. Related tables are presented in Appendix E.

4.3 LCE's Obtained at the Largest T_v of the Corresponding Choices versus the Vector Size m for $p = 3$

It is also of interest to find the variation of LCE values while the vector size m is increasing. In order to investigate this subject, we fix the field size p to a small value 3, to diminish the computational cost so that the vector size m can be increased as much as possible. For each m, we exhaustively produce all possible sequences with the largest vector period T_v, and then use the Berlekamp-Massey algorithm to compute the linear complexities of all sequences in this set. We have been able to increase the vector size m up to 7, 5, 13, 9 and 9 for Choices 1, 2, 3a, 3b and 3c, respectively. The differences in the attained upper limits of m are related to the complexity of corresponding choice of polynomial iterations, Choice 3a being the simplest of all, it can be used with the highest m for evaluation of the LCE's within the set of largest period sequences found exhaustively. Corresponding values of minimum, average and maximum LCE's are drawn in Figure 4.2 for the five polynomial choices and the computation times are given in Appendix A.

Figure 4.2 (a) indicates that the LCE values of the largest period sequences generated by Choice 1 in $\mathbb{F}_3$ are at most 0.67 for odd m, but they can reach the highest point (LCE ≥ 0.95) when m is even. However, it is still not possible to recommend the largest period sequences of Choice 1 as sufficiently random sequences, because of the minimum LCE's of the set that may be as small as 0.6. Another observation is from Figure 4.2 (b), showing that LCE values of the largest period sequences generated by Choice 2 in $\mathbb{F}_3$ are 0.67 that seems independent of m. On the other hand, Figure 4.2 (c) points out that LCE values of the largest period sequences generated by Choice 3a in $\mathbb{F}_3$ can be higher when m is a power of $p = 3$. However, they are decreasing dramatically as m grows; and relatively high values at $m = 3$ or 9 are not sufficient to declare this set of sequences as a reliable source of pseudo-random noise generator either; especially considering their LCE minima that remains around 0.2 independently of m. Figures 4.2 (d) and (e) designate that the LCE performances of the largest period sequences of Choices 3b and 3c in $\mathbb{F}_3$ are quite similar, they seem better than that of Choice 3a; however, because of their minimum LCE curves around 0.4, they are also not recommendable as pseudo-random sequences. Related tables are presented in Appendix E.

Overall comparison of the LCE performances in $\mathbb{F}_3$, among the largest period sequences of the five polynomial choices given in Table 1.1 is in favor of Choice 1; however one can still not say that one of the largest period Choice 1 sequences chosen at random has sufficiently high LCE, since it may be as low as 0.6.

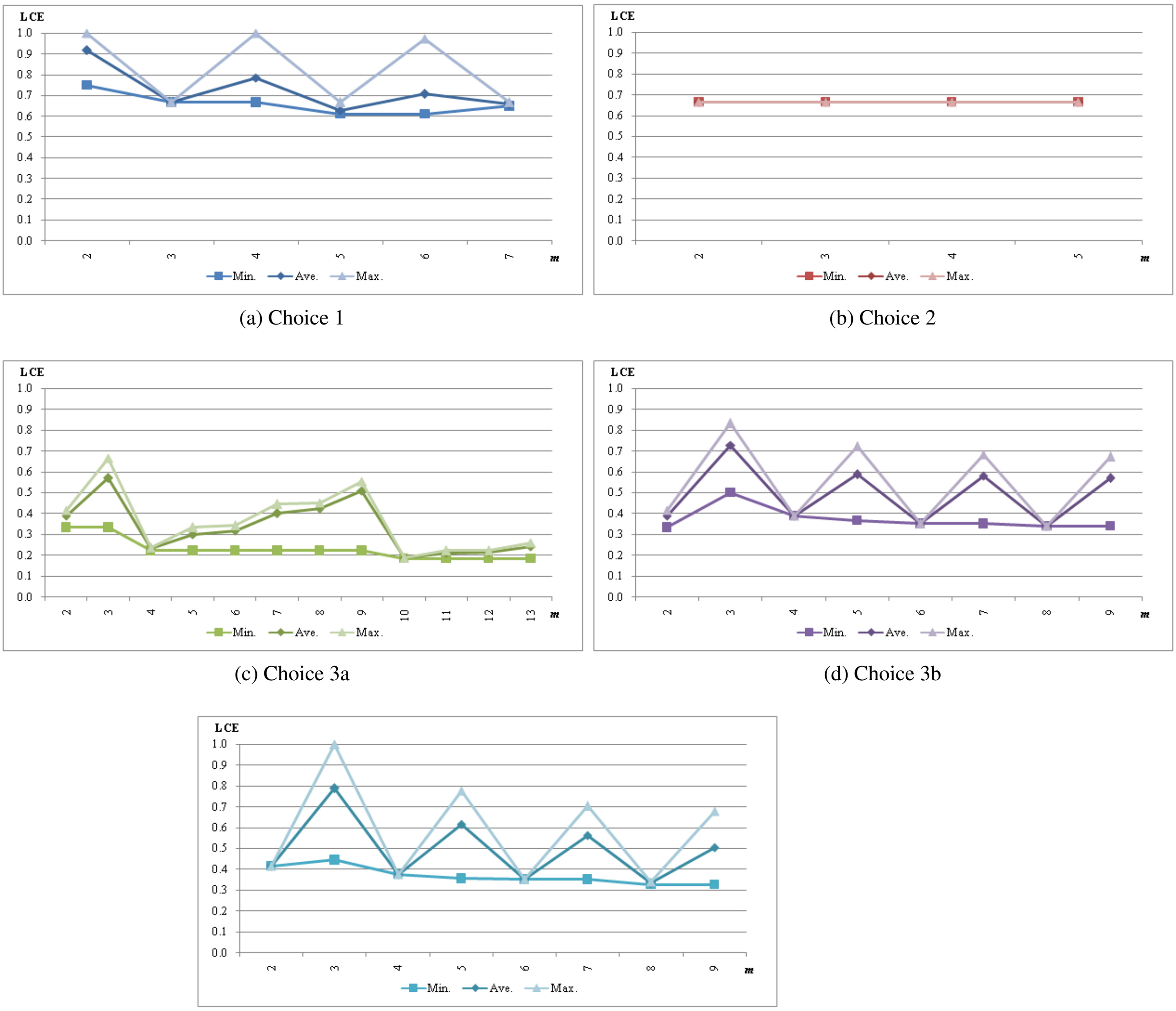


(a) Minimum

(b) Average

(c) Maximum

Figure 4.1: Variation of the: (a) minimum, (b) average and (c) maximum LCE's obtained at the largest T_v of the corresponding choice versus p, at $m = 2$



(a) Choice 1

(b) Choice 2

(c) Choice 3a

(d) Choice 3b

(e) Choice 3c

Figure 4.2: Minimum, average and maximum LCE obtained at the largest T_v of: (a) Choice 1, (b) Choice 2, (c) Choice 3a, (d) Choice 3b and (e) Choice 3c, versus m for $p = 3$

4.4 Linear Complexity of Maximum-Period Sequences

In this section, we investigate the randomness of maximum-period sequences generated by Choice 4, which is the last polynomial choice [20] shown in Table 1.1. Our measure of randomness is the linear complexity computed by the Berlekamp-Massey algorithm.

4.4.1 Generating Maximal Period Sequences with Choice 4

In the general description given by (1.3) of multivariate polynomial iterations,

$$\begin{gathered}
F_1(\mathbf{X}) = X_1G_1(X_2,\ldots,X_m) + H_1(X_2,\ldots,X_m),\\
F_2(\mathbf{X}) = X_2G_2(X_3,\ldots,X_m) + H_2(X_3,\ldots,X_m),\\
\vdots\\
F_{m-1}(\mathbf{X}) = X_{m-1}G_{m-1}(X_m) + H_{m-1}(X_m),\\
F_m(\mathbf{X}) = g_mX_m + h_m,
\end{gathered}$$

Choice 4 generates the maximum-period sequences that have the maximum period efficiency, PE $= T/p^m = 1$, by substituting

$$G_i(X_{i+1},\ldots,X_m) = 1,\;\; g_m = 1$$

and

$$H_i(X_{i+1},\ldots,X_m) = X_{i+1}^{p-1}\;\ldots\;X_m^{p-1},\;\; h_m \neq 0$$

for

$$i = 1,\ldots,m-1.$$

Hence, the iterations of Choice 4 sequences are given by

$$\begin{gathered}
F_1(X) = X_1 + X_2^{p-1}\ldots X_m^{p-1},\\
F_2(X) = X_2 + X_3^{p-1}\ldots X_m^{p-1},\\
\vdots\\
F_{m-1}(X) = X_{m-1} + X_m^{p-1},\\
F_m(X) = X_m + h_m.
\end{gathered}\tag{4.1}$$

4.4.2 Linear Complexity Efficiency Values of Choice 4 Sequences

LCE (L/T) values of the maximum-period sequences generated by Choice 4 are computed and tabulated in Table 4.1 for field sizes $3 \leq p \leq 31$ and vector sizes $2 \leq m \leq 7$ found over all possible initial values $\mathbf{X} = (X_1, \ldots, X_m)$, and h_m (number of all possible initial values is equal to $p^m(p-1)$).

One can observe from Table 4.1 that the LCE values of the sequences generated by Choice 4 are very poor and equal to the fixed value given in Proposition 4.1.

Proposition 4.1. *Linear complexity L of a sequence generated by Choice 4 for given m and p, is equal to $m(p^{m-1}+1)$ within the set $2 < p \leq 31$ considered in this work.*

On the other hand, the efficiency decreases while p and m are increasing, since $L/T = m(p^{m-1}+1)/mp^m = (p^{-1}+p^{-m})$. For $p = 2$, LCE $= L/T$ is worse than $(p^{-1} + p^{-m})$, and it is approximately equal to p^{-1} as the linear complexities given in Table 4.2 indicate.

Moreover, the minimal polynomials of these sequences found by using the Berlekamp-Massey algorithm are observed to obey the general form given in Proposition 4.2.

Proposition 4.2. *Minimal polynomials of Choice 4 sequences for given m, p and L, are in the form $c(x) = 1 + (p-1)x^m + (p-1)x^{L-m} + x^L$ in terms of m, p and L within the set $2 < p \leq 31$ considered in this work.*

4.4.3 Example Sequences Generated by Choice 4 with Poor Randomness

Although Choice 4 is proposed as a method to obtain maximum length sequences with vector period p^m, one can produce simple examples for a field size $p = 2$ with poor randomness properties using Choice 4. In the general description given by Equation (1.3) of multivariate polynomial iterations,

$$F_1(\mathbf{X}) = X_1 G_1(X_2, \ldots, X_m) + H_1(X_2, \ldots, X_m),$$

$$\vdots$$

$$F_{m-1}(\mathbf{X}) = X_{m-1} G_{m-1}(X_m) + H_{m-1}(X_m),$$

$$F_m(\mathbf{X}) = g_m X_m + h_m,$$

with

$$G_i \,, H_i \in \mathbb{F}_p[X_{i+1}, \ldots, X_m],\ i = 1, \ldots, m-1$$

and

$$g_m,\ h_m \in \mathbb{F}_p,\ g_m \neq 0. \tag{3}$$

Table 4.1: Linear complexity efficiency values of the maximum-period sequences

p	m	Count	T_v	T	L	LCE
3	5	486	243	1215	410	0.34
3	6	1458	729	4374	1464	0.33
3	7	4374	2187	15309	5110	0.33
5	3	500	125	375	78	0.21
5	4	2500	625	2500	504	0.20
5	5	12500	3125	15625	3130	0.20
7	3	2058	343	1029	150	0.15
7	4	14406	2401	9604	1376	0.14
11	2	1210	121	242	24	0.10
11	3	13310	1331	3993	366	0.09
13	2	2028	169	338	28	0.08
13	3	26364	2197	6591	510	0.08
17	2	4624	289	578	36	0.06
17	3	78608	4913	14739	870	0.06
19	2	6498	361	722	40	0.06
19	3	123462	6859	20577	1086	0.05
23	2	11638	529	1058	48	0.05
23	3	267674	12167	36501	1590	0.04
29	2	23548	841	1682	60	0.04
29	3	682892	24389	73167	2526	0.03
31	2	28830	961	1922	64	0.03
31	3	893730	29791	89373	2886	0.03

Table 4.2: Linear complexity efficiency values of the maximum-period sequences for $p = 2$

m	Count	T_v	T	L	LCE
2	4	4	8	4	0.500
3	8	8	24	12	0.500
4	16	16	64	32	0.500
5	32	32	160	82	0.513
6	64	64	384	190	0.495
7	128	128	896	448	0.500
8	256	256	2048	1024	0.500
9	512	512	4608	2304	0.500
10	1024	1024	10240	5120	0.500
11	2048	2048	22528	11265	0.500
12	4096	4096	49152	243	0.500

Table 4.3: An example of the sequences generated by Choice 4

p	m	$\mathbf{F}^{(0)}, \ldots, \mathbf{F}^{(T_v)}$
2	2	(0 0), (0 1), (1 0), (1 1)
2	3	(0 0 0), (0 0 1), (0 1 0), (0 1 1), (1 0 0), (1 0 1), (1 1 0), (1 1 1)

Choice 4 is obtained by substituting $G_i(X_{i+1}, \ldots, X_m) = 1$, $gm = 1$ and $H_i(X_{i+1}, \ldots, X_m) = X_{i+1}^{p-1} \ldots X_m^{p-1}$, $h_m \neq 0$.

In a finite field with 2 elements, $H_i(X_{i+1}, \ldots, X_m) = X_{i+1}^{p-1} \ldots X_m^{p-1} = X_{i+1} \ldots X_m$, hence for $p = 2$ and $h_m = 1$, polynomial iterations with Choice 4 are reduced to:

$F_1(\mathbf{X}) = X_1 + X_2 \ldots X_m, \ldots, F_{m-1}(\mathbf{X}) = X_{m-1} + X_m$, $F_m(\mathbf{X}) = X_m + 1$;

more specifically for $m = 2$ to $F_1(\mathbf{X}) = X_1 + X_2$, $F_2(\mathbf{X}) = X_2 + 1$, and for $m = 3$ to $F_1(\mathbf{X}) = X_1 + X_2X_3, \ldots, F_2(\mathbf{X}) = X_2 + X_3$, $F_3(\mathbf{X}) = X_3 + 1$.

Now taking the initial state X as the all-zero vector, the sequences given in Table 4.3 are generated. Vector sequences in Table 4.3 clearly exhibit a non-random behavior, since they are ordered lexicographically.

4.5 Conclusion

Ostafe's multivariate iterations (1.3) ([18]-[25]) can be used with different polynomial choices (see Table 1.1). For all polynomial choices considered in this work (and in the literature that we have encountered), we have evaluated the performance of the largest period sequences exhaustively, in terms of the efficiency of linear complexity, as computed by the Berlekamp-Massey algorithm. The result is not encouraging, because the best of all choices seems to be Choice 1 that is still not good enough for recommendation as a PRNG.

Choice 4 seems to be the least random choice. Because, for the sequences generated by Choice 4 [20] that are known to have the maximum period efficiency PE $= 1$, we obtain extremely low linear complexity values. Their linear complexity efficiency is equal to $(p^{-1} + p^{-m})$ for $p > 2$; hence it decreases with increasing p and m. For $p = 2$, the LCE of a sequence generated by Choice 4 is worse than $(p^{-1} + p^{-m})$, and it is approximately equal to p^{-1}.

CHAPTER 5

LINEAR COMPLEXITY ANALYSIS OF SEQUENCES GENERATED WITH RANDOM INITIAL VALUES

5.1 Introduction

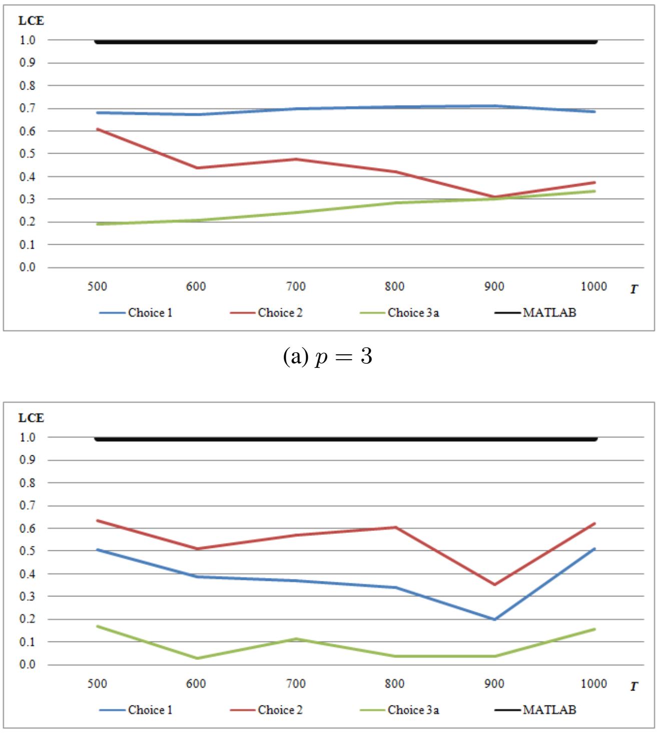
In the previous chapters, periods and linear complexities of Ostafe's multivariate polynomial iterations (1.3) have been examined exhaustively over all possible initial values. In this chapter, we investigate the linear complexities of these sequences when the initial values are chosen randomly as in cryptographic applications. We only consider Choice 1, 2 and 3a. Choice 3b and 3c are not included because of their resemblance to Choice 3a. Choice 4 [20] is not included for two reasons: *i*) the period of Choice 4 sequences is fixed as $T = mp^m$, which is not a flexible value, *ii*) Choice 4 is known to have the poorest LCE among other choices as the results of Section 4.4 indicate.

In Section 5.2, we consider Choice 1, 2, 3a and MATLAB's randi(.) function [1], and for each choice we compute the average LCE values of 100 sequences having similar periods. In Section 5.3 and 5.4, we investigate the effect of the field size p and the number of polynomials m respectively on the LCE values, again over the sets with 100 similar-period sequences. In Section 5.5, we omit the constraint of similar periods and consider 100 sequences with variable periods. Then we analyze the effect of the field size p and the number of polynomials m on the linear complexity.

5.2 Randomness Comparison with MATLAB's randi(.) Sequence

In order to investigate the randomness properties of sequences generated by (1.3), we consider fixed length sequences within the range $T \pm \alpha T$ produced by Choice 1 ([21]), 2 ([18]) and 3a ([19]) (with our choice of $H_i = X_i + 1$ shown in Table 1.1), as well as the reference method, MATLAB's **randi(.)** function.

[1] MATLAB (Matrix Laboratory) is a multi-paradigm numerical computing environment, widely used in mathematics and engineering; and enables the user to execute mathematical operations easily and efficiently. It has used George Marsaglia's Ziggurat algorithm [13] developed by George Marsaglia of Florida State University in order to produce pseudo-random numbers [15].



(a) $p = 3$

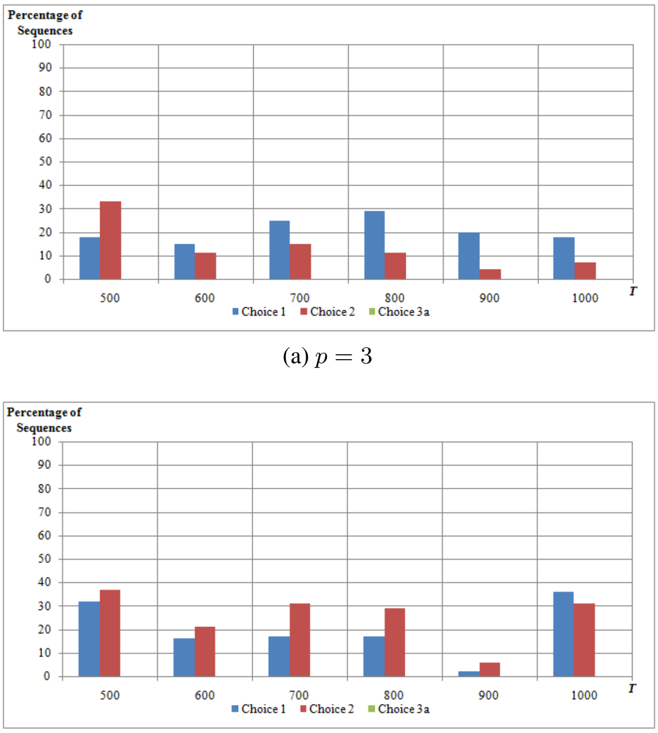
(b) $p = 13$

Figure 5.1: Randomness comparison of Ostafe's and MATLAB's **randi(.)** sequences for (a) $p = 3$, (b) $p = 13$, in terms of the LCE as defined in this work by "linear complexity L divided by the period T of the sequence" (For each T, the average LCE is computed over 100 sequences.)

For each polynomial choice and the field characteristic $3 \leq p \leq 13$, we pick up a period T from the set $T \in \{500, 600, 700, 800, 900, 1000\}$ (and from the set $T \in \{500, 800, 1000\}$ for $17 \leq p \leq 31$ as given in Appendix D) and generate 100 sequences having periods within the range $T \pm 0.05T$, by assigning random values to the remaining parameters (i.e., the number of polynomials m, $\mathbf{X} = (X_1, \ldots, X_m)$, $\mathbf{G} = (G_1, \ldots, G_{m-1})$, $\mathbf{H} = (H_1, \ldots, H_{m-1})$, a_i, g_m and h_m) of the related polynomial choice. We compute the linear complexity L of each sequence, and find how close it is to the period by computing its Linear Complexity Efficiency (LCE), L/T. After producing 100 such sequences for a given p and T, compute the average linear complexity efficiency (L/T) over 100 sequences in each period group.

In Appendix F we repeat this experiment for a wider range of field sizes as well; i.e., $17 \leq p \leq 31$ and periods $T \in 500, 800, 1000$. Figure 5.1(a) and (b) depict the average LCE values for $p = 3$ and 13, found over 100 sequences at each T. The results for other field sizes $5 \leq p \leq 11$ are also given in Appendix F. We also draw the linear complexity profiles (see Definition 3.3) of the generated sequences in Appendix J.

One can observe from Figure 5.1 that the average LCE values of the reference sequences generated by MATLAB look perfectly random, with average LCE almost equal to 1. On the other hand, Ostafe's sequences, Choice 1 and 2, have lower LCE values for all T's in the given set. Randomness properties of Choice 3a is the worst among the first three choices of Table 1.1, since its average LCE does not exceed 0.4.



(a) $p = 3$

(b) $p = 13$

Figure 5.2: Percentage of sequences with LCE ≥ 0.95 for (a) $p = 3$, (b) $p = 13$, using three choices

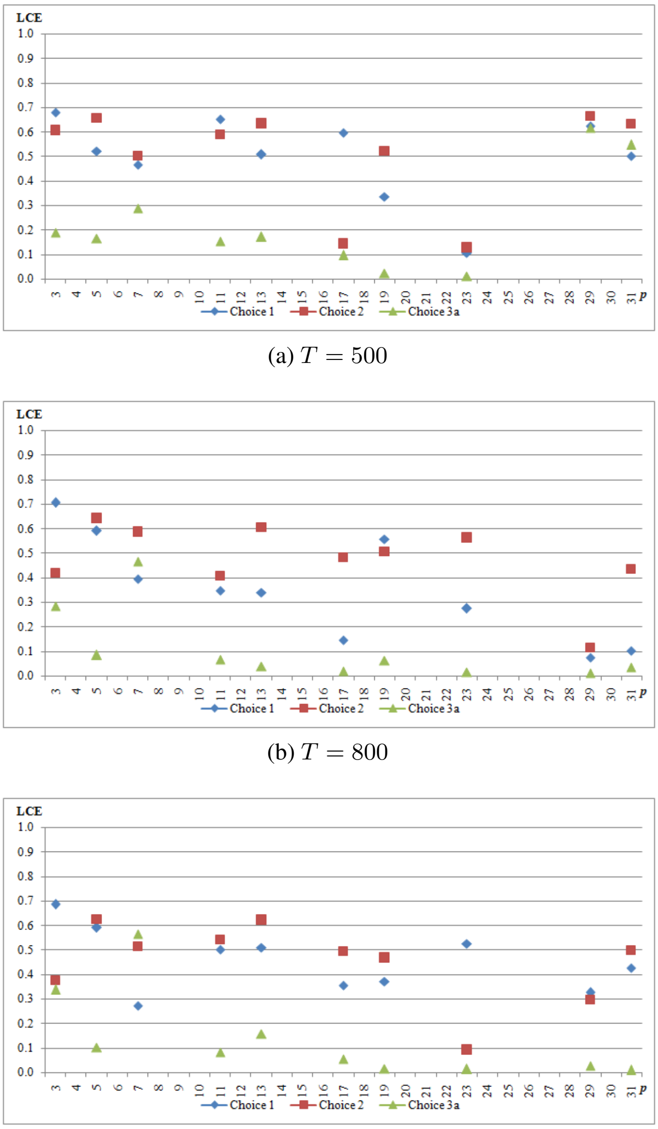
Although the average LCE values of all choices given in Table 1.1 are small, some of these random sequences may seldomly yield high LCE's as well. In Figure 5.2, we plot the percentage of sequences with LCE ≥ 0.95 for $p = 3$ and 13, over the generated 100 sequences.

Figure 5.2 shows that for field sizes $p = 3$ and 13, Choice 1 and 2 generate highly random sequences with low percentages (not reaching 40%). On the other hand, Choice 3a is not able to generate them at all for both $p = 3$ and 13. The results for other field sizes $5 \leq p \leq 11$ given in Figure F.2 give similar information.

5.3 Effect of the Field Size p for Similar Period Sequences

In order to understand whether the low LCE values of Choice 1, 2 and 3a sequences can be improved by increasing the field characteristic p, experiments are performed for each polynomial choice by averaging over a set of 100 sequences that have the periods $T \pm 0.05T$, where $T = 500, 800$ or 1000. Average LCE values of the 100 sequences generated by the three polynomial choices are sketched in Figure 5.3 for $3 \leq p \leq 31$.

It is observed from Figure 5.3 that increasing the field size p does not improve the LCE values of the random-initial-value sequences generated by Choice 1, 2 and 3a for the given T and p sets.



(a) $T = 500$

(b) $T = 800$

(c) $T = 1000$

Figure 5.3: LCE values for a sequence of length (a) $T = 500$, (b) $T = 800$, (c) $T = 1000$, in variable field sizes, $3 \leq p \leq 31$ using the three choices given in Table 1.1

Table 5.1: LCE values of the sequences generated by Choice 1 for $p = 3$, $T = 1000$, $m = 9$

$(X_1, \ldots, X_9)$	$(h_1, \ldots, h_9)$	g_9	LCE
(2,2,2,0,0,1,1,1,1)	(0,0,2,0,1,1,1,1,1)	1	0.19
(0,2,0,2,1,1,1,1,2)	(2,1,1,0,1,1,0,0,0)	1	0.58
(0,2,1,2,0,2,1,1,0)	(1,0,0,1,1,0,0,2,1)	1	0.61
(0,1,1,1,1,2,2,1,0)	(1,1,1,0,2,1,0,2,1)	2	0.62
(2,0,1,1,1,1,2,2,2)	(1,1,1,1,0,2,1,0,0)	2	0.63
(2,2,1,2,1,1,1,0,1)	(1,0,2,1,0,0,2,1,0)	1	0.64
(0,1,0,2,1,2,2,2,0)	(2,2,1,2,1,0,0,2,0)	1	0.65
(2,1,1,0,1,2,1,2,2)	(2,1,2,1,2,1,0,0,1)	2	0.66
(0,1,1,1,2,1,1,2,1)	(0,1,1,1,0,0,2,1,0)	1	0.67
(2,0,2,1,2,1,1,1,1)	(1,1,0,0,2,1,1,1,0)	1	0.68
(2,1,1,1,0,1,2,1,2)	(0,1,1,2,1,2,1,0,0)	2	0.83
(1,1,1,2,2,1,2,1,1)	(1,2,2,1,2,1,0,1,1)	2	0.88
(0,1,2,1,2,1,2,1,1)	(2,2,2,1,0,0,2,1,2)	2	0.90
(0,1,0,2,2,1,2,1,1)	(0,1,2,0,0,2,1,1,1)	1	0.98
(0,1,1,0,2,2,2,2,1)	(1,0,2,1,1,1,1,1,0)	1	0.99
(0,1,2,1,0,2,0,2,2)	(1,2,0,2,1,1,1,1,1)	1	1.00

5.4 Effect of the Number of Polynomials m for Similar Period Sequences

The topic of this section is to see the effect of the number of polynomials m on the LCE values of the random-initial-value sequences produced by the three choices given in Table 1.1. In order to investigate the effect of m for the first three choices in Table 1.1, we generate ternary $(p = 3)$ sequences of period $T = 1000 \pm 50$ and tabulate the LCE values of the ones having the same m. For instance, LCE values of Choice 1 sequences generated with $m = 9$ polynomials are given in Table 5.1. The fact that the LCE values for the same m occupy a wide range between 0.19 and 1, indicates the existence of very little correlation between these two parameters, namely m and (LCE $= L/T$).

Similarly, Table 5.2 presented for Choice 2 proves that if one desires a sequence of length $T = 1000{\pm}50$, $m = 12$ polynomials may generate a ternary sequence having an LCE between 0.66 and 1. As for Choice 3a, Table 5.3 shows that $m = 19$ polynomials iterated according to Equation (1.3) produce ternary sequences of length $T = 1000{\pm}50$ with LCE values ranging between 0.30 and 0.37.

Table 5.2: LCE values of the sequences generated by Choice 2 for $p = 3$, $T = 1000$, $m = 12$

$(X_1,\ \ldots,\ X_{12})$	$(h_1,\ \ldots,\ h_{12})$	$(a_1,\ \ldots,\ a_{12})$	g_{12}	LCE
(2,2,1,2,0,1,0,0,2,1,1,1)	(2,2,1,1,0,1,1,0,1,1,1,1)	(2,2,2,2,2,2,2,2,2,2,2,2)	1	0.66
(1,0,2,2,0,1,0,0,1,1,0,0)	(0,0,1,0,0,0,0,2,2,2,2,0)	(2,2,2,2,2,2,2,2,2,2,2,2)	2	0.67
(1,0,0,1,2,0,1,2,1,2,1,0)	(1,0,0,1,2,0,2,0,2,1,2,1)	(2,2,2,2,2,2,2,2,2,2,2,2)	1	1.00

Table 5.3: LCE values of the sequences generated by Choice 3a for $p = 3$, $T = 1000$, $m = 19$

$(X_1,\ \ldots,\ X_{19})$	$(g_1,\ \ldots,\ g_{19})$	h_{19}	LCE
(0,2,1,0,1,2,2,2,2,0,2,1,2,0,2,2,1,2,1)	(2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2)	2	0.30
(1,2,0,2,1,0,0,2,2,1,0,2,1,1,2,0,1,0,0)	(2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2)	0	0.31
(2,1,2,0,1,2,1,2,1,2,0,0,0,1,0,0,0,2,1)	(2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2)	2	0.33
(0,2,1,0,0,0,1,0,2,0,0,1,2,2,2,0,0,1,2)	(2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2)	0	0.35
(1,0,0,1,1,1,0,1,2,1,2,0,0,2,0,2,2,1,0)	(2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2)	1	0.37

Table 5.4 shows that 42% of the generated Choice 3a sequences with $T = 1000\pm 50$ use 19 polynomials, and the remaining 58% is produced with 18 polynomials; all resulting in poor LCE values varying in a narrow range of $[0.26, 0.37]$. The slight increase in the average LCE values corresponding to the slight increase in the number of polynomials from 18 to 19 is not a sufficient indicator to claim some correlation between m and LCE parameters.

It should be noted that the data in Table 5.4 is the instance of the general experiment and the remaining data given in Table H.1 to H.9 has similar characteristics; i.e., LCE values of the sequences generated by Choice 3a vary in a narrow range.

5.5 Effect of p and m for Variable Period Sequences

In the previous two sections, we have computed average LCE's over 100 sequences with similar periods. In order to see the effect of the field size p and the vector size m within the sets of sequences having variable periods, we perform one more analysis: we generate 100 sequences with random initialization for three choices, p and m where $3 \leq p \leq 43$ and $2 \leq m \leq 4$, without fixing the periods and we compute the average

Table 5.4: LCE values of the sequences generated by Choice 3a for $p = 3$, $T = 1000$

m	Minimum	Maximum	Average	Count
18	0.26	0.33	0.32	58
19	0.30	0.37	0.36	42

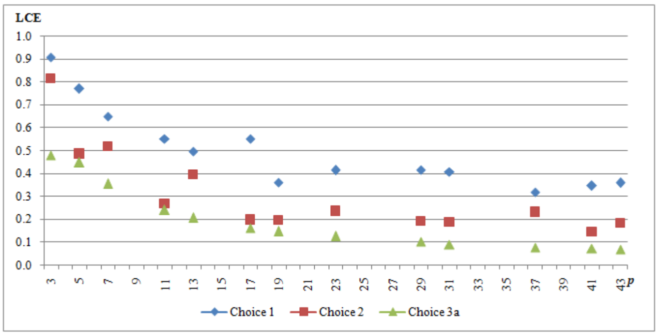
LCE in each set. For avoiding the shortest-period sequences which are practically meaningless, we only consider vector periods $T_v \geq p-1$. In each set of 100 sequences, percentages of those with LCE ≥ 0.95 and corresponding period efficiencies are given in Appendix G, where one can see that a sequence with high LCE and PE is not possible to find.

In Figure 5.4, we depict the average LCE's of the mentioned three choices versus the field size p for three different vector sizes m. We observe no clear dependence on m, but the average LCE values decrease with increasing p in general. One may notice that for $m > 2$ the average LCE's of Choice 1 sequences seem a little more promising; i.e., for $m = 3$ average LCE remains around 0.7 and for $m = 4$ it does not fall below 0.75 with increasing p. However, in generating random sequences it is necessary to obtain a high-LCE sequence in each trial, so we also draw the minimum LCE values within the same 100-element sets in Figure 5.5. Minimum LCE's clearly indicate that Choice 1 is also not preferable as a PRNG, since it may produce sequences having LCE's as small as those of the other choices.

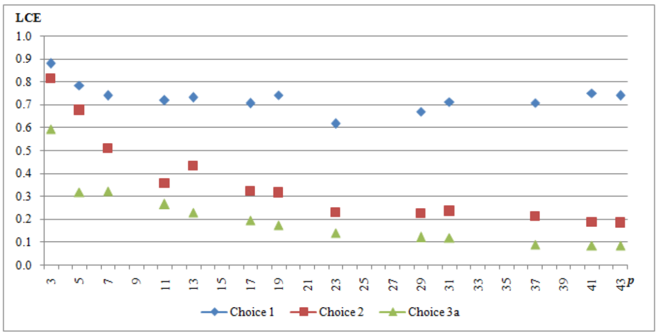
5.6 Conclusion

We have compared the randomness of the sequences generated with random initialization by Choice 1, 2 and 3a of Ostafe's polynomial iterations (1.3), as well as the reference method MATLAB's **randi(.)** function, in terms of their linear complexities. Within each set of 100 sequences produced by random initialization, we have observed that none of the Ostafe's polynomial choices are able to produce sequences as random as those generated by MATLAB's **randi(.)** function. With very low probabilities, Choice 1 and 2 can generate sequences having high linear complexities but the corresponding period efficiencies are quite low. On the other hand, Choice 3a does not generate any sequence with high linear complexity $(L/T \geq 0.95)$ at all.

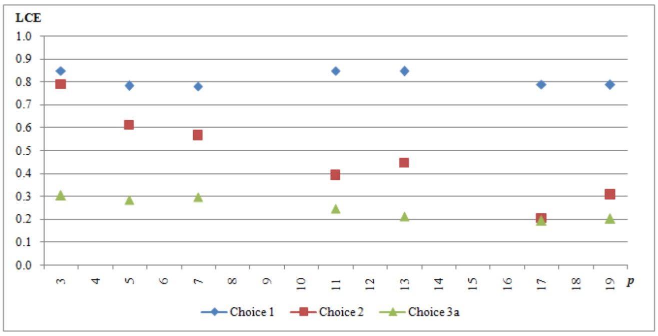
The number of polynomials; i.e., the vector size m has no noticeable effect on the linear complexities of Ostafe's sequences. The other observation is that increasing the field size, p, decreases the minimum linear complexity efficiencies for all three choices.



(a) $m = 2$

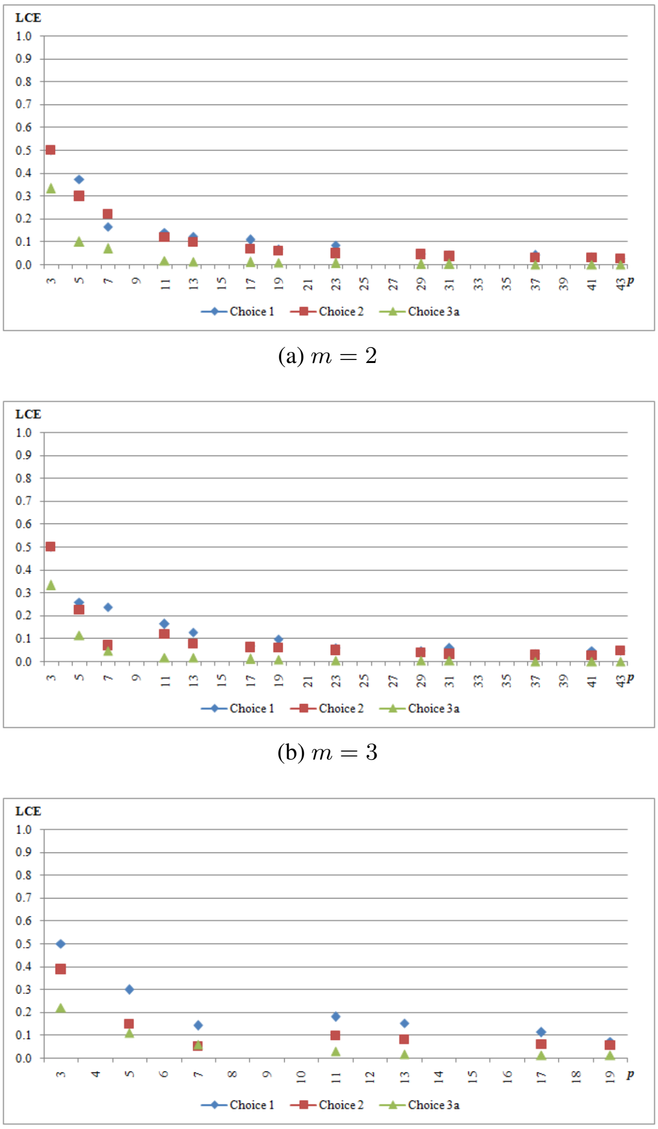


(b) $m = 3$



(c) $m = 4$

Figure 5.4: Variation of the average LCE's of the sequences generated by random initialization of the corresponding choice versus p at: (a) $m = 2$, (b) $m = 3$, (c) $m = 4$



(a) $m = 2$

(b) $m = 3$

(c) $m = 4$

Figure 5.5: Variation of the minimum LCE's of the sequences generated by random initialization of the corresponding choice versus p at: (a) $m = 2$, (b) $m = 3$, (c) $m = 4$

CHAPTER 6

CONCLUSION

In this study, we analyze the randomness properties of the scalar sequences of length T, obtained from the vector sequences of Ostafe and Shparlinski ([18]-[25]) generated in $\mathbb{F}_p$ by the m-variate m-polynomial recursive method (1.3). Our analysis depends on two basic approaches: the period and linear complexity distributions of the produced sequences. In order to measure the potential of Ostafe's polynomial iterations as a candidate for a PRNG, we define two parameters; namely, the "period efficiency (PE)" and the "linear complexity efficiency (LCE)". These parameters are computed by normalizing the period and the linear complexity with respect to their maximum possible values, T and mp^m respectively; hence, they both take values in the interval [0,1].

Firstly, we have performed an exhaustive search in order to find the distribution of the periods generated by the five suggested choices: Choice 1 [21], Choice 2 [18] and Choice 3a, 3b, 3c [19]. Our exhaustive search for the distribution of the vector period T_v, in the fields of size $p = 3$ with $2 \leq m \leq 5$, $p = 5, 7$ with $m = 2, 3$, and $p = 11, 13$ with $m = 2$, shows that there is no maximum-period sequence for Choice 1, 3a, 3b and 3c. Only Choice 2 can generate maximum-period sequences with vector period p^m, however their existence probability is less than 3$\%$ if $p > 3$.

Secondly, we have executed an exhaustive search for investigating the linear complexities of the sequences, as computed by the Berlekamp-Massey algorithm. We observe that Choice 1 generates sequences with LCE's ≥ 0.95 more efficiently than Choices 2, 3a and 3b. Still, the percentage of Choice 1 sequences with LCE ≥ 0.95 is less than 48$\%$ and corresponding PE's are less than 0.44. We also notice that all five polynomial choices produce less sequences with LCE ≥ 0.95 at high vector periods than at low vector periods.

Thirdly, we have evaluated the LCE performance of the largest period sequences exhaustively. Similar to the previous results, Choice 1 seems to be the best in terms of the LCE; however, the corresponding PE's are not satisfactory. On the other hand, Choice 4 [20], which is known to have the maximum period efficiency PE $= 1$, has extremely low linear complexity values. The corresponding linear complexity efficiency is equal to $(p^{-1} + p^{-m})$ for $p > 2$; hence it decreases with increasing p and m. For $p = 2$, the LCE of a sequence generated by Choice 4 is worse than $(p^{-1} + p^{-m})$, and it is approximately equal to p^{-1}.

Finally, getting rid of the exhaustive search and using random initialization instead; we have been able to increase the values of p and m, and compared the randomness of Ostafe's sequences with the reference method, MATLAB's **randi(.)** function. We have observed that none of the Ostafe's polynomial choices are able to produce sequences as random as those generated by MATLAB's **randi(.)** function. The number of polynomials; i.e., the vector size m has no noticeable effect on the linear complexities of Ostafe's sequences. The other observation is that increasing the field size, p, decreases the minimum linear complexity efficiency in general. We have also seen that with very low probabilities, Choice 1 and 2 can generate sequences having high linear complexities. However, these results are not encouraging enough to propose any of Ostafe and Shparlinski's choices as a PRNG.

As a result of this study, one can say that the sequences generated by the multivariate polynomial iterations method with six mentioned choices do not satisfy the desired randomness properties, with respect to the period and the linear complexity.

REFERENCES

[1] Berlekamp, E. R. *Algebraic Coding Theory*. McGraw-Hill, NY, 1968.

[2] Boyar, J. Inferring a sequence generated by a linear congruence. *Proceedings of the 23rd Annual IEEE Symposium on the Foundations of Computer Science* (1982), 153–159.

[3] Boyar, J. Inferring sequences produced by pseudo-random number generators. *Journal of the ACM 36* (1989), 129–141.

[4] Caelli, W., Dawson, E., Nielsen, L., and Gustafson, H. Crypt-x statistical package manual, measuring the strength of stream and block ciphers.

[5] Eichenauer, J., Grothe, H., and Lehn, J. Marsaglia's lattice test and nonlinear congruential pseudo random number generators. *Metrika 35* (1988).

[6] Glen, A. On the period length of pseudorandom number sequences, honours thesis.

[7] Knuth, D. E. *The art of computer programming, 2nd ed.* Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1981.

[8] Knuth, D. E. Deciphering a linear congruential encryption. *IEEE Trans. Inform. Theory 31* (1985), 49–52.

[9] Krawczyk, H. How to predict congruential generators. *Journal of Algorithms 13* (1992), 527–545.

[10] L'Ecuyer, P., and Simard, R. Testu01: A c library for empirical testing of random number generators. *ACM Trans. Math. Softw. 33* (2007).

[11] Lehmer, D. H. Mathematical methods in large-scale computing units. *Harvard University Press, Cambridge, Mass. 180* (1951), 141–146.

[12] Marsaglia, G. The marsaglia random number cdrom including the diehard battery of tests of randomness.

[13] Marsaglia, G., and Tsang, W. W. The ziggurat method for generating random variables. *Journal of Statistical Software 5*, 8 (2007).

[14] Massey, J. L. Shift-register synthesis and bch decoding. *IEEE Trans. Info. Theory 15*, 1 (1969), 122–127.

[15] Moler, C. The ziggurat random normal generator, mathworks. `http://blogs.mathworks.com/cleve/2015/05/18/the-ziggurat-random-normal-generator`, 2015. [Online; accessed January 2016].

[16] NIEDERREITER, H. *Random number generation and Quasi-Monte Carlo methods.* SIAM Press, 1992.

[17] NIEDERREITER, H. Linear complexity and related complexity measures for sequences. *Lecture Notes in Computer Science Volume 2904* (2003), 1–17.

[18] OSTAFE, A. Multivariate permutation polynomial systems and pseudorandom number generators. *Finite Fields Appl. 16* (2010), 144–154.

[19] OSTAFE, A. Pseudorandom vector sequences derived from triangular polynomial systems with constant multipliers. *in: Lecture Notes in Comput. Sci., Springer-Verlag, Berlin 79* (2010), 62–72.

[20] OSTAFE, A. Pseudorandom vector sequences of maximal period generated by triangular polynomial dynamical systems. *Des. Codes Cryptogr. 63* (2012), 59–72.

[21] OSTAFE, A., AND SHPARLINSKI, I. On the degree growth in some polynomial dynamical systems and nonlinear pseudorandom number generators. *Math. Comp. 79* (2010), 501–511.

[22] OSTAFE, A., AND SHPARLINSKI, I. Pseudorandom numbers and hash functions from iterations of multivariate polynomials. *Cryptogr. Commun. 2* (2010), 49–67.

[23] OSTAFE, A., AND SHPARLINSKI, I. Degree growth, linear independence and periods of a class of rational dynamical systems. *Arithmetic, geometry, cryptography and coding theory, Contemp. Math., Amer. Math. Soc., Providence, RI 574* (2012), 131–143.

[24] OSTAFE, A., AND SHPARLINSKI, I. On the power generator and its multivariate analogue. *J. Complexity 28*, 2 (2012), 238–249.

[25] OSTAFE, A., SHPARLINSKI, I., AND WINTERHOF, A. On the generalized joint linear complexity profile of a class of nonlinear pseudorandom multisequences. *Adv. Math. Comm. 4 3* (2010), 369–379.

[26] OSTAFE, A., AND SHPARLINSKI, I.E., P. E. On pseudorandom numbers from multivariate polynomial systems. *Finite Fields and Their Appl. 6* (2010), 320–328.

[27] RITTER, T. The efficient generation of cryptographic confusion sequences. *Cryptologia 15* (1991), 81–139.

[28] RUEPPEL, R. A. *Analysis and Design of Stream Ciphers.* Springer-Verlag, Berlin, 1986.

[29] RUKHIN, A., SOTO, J., NECHVATAL, J., BARKER, E., LEIGH, S., LEVENSON, M., BANKS, D., HECKERT, A., DRAY, J., AND VO, S. Statistical test suite for random and pseudorandom number generators for cryptographic applications. *NIST Special Publication* (2001).

[30] SHANNON, C. A mathematical theory of communication. *The Bell Systems Technical Journal 27* (1948), 370–423 and 623–656.

APPENDIX A

COMPUTATION TIMES OF THE EXHAUSTIVE SEARCHES FOR FIVE CHOICES

In this appendix, we aim to present the dependence of the required computation times on the parameters p and m. Table A.1 shows the computation times of the exhaustively calculated period efficiency (PE) and linear complexity efficiency (LCE) of the sequences generated by five choices for some of the considered cases. All computations are performed via Intel(R) Xeon(R) CPU 3.70 GHz.

As can be seen from A.1, Choice 2 is the most time consuming one in all polynomial choices.

Table A.1: Computation times (in seconds) of the exhaustive PE and LCE searches for five choices

p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
3	2	0.26	0.29	0.11	0.13	0.17
3	3	2.00	3.48	0.13	0.17	0.21
3	4	29.02	114.70	0.63	1.02	0.94
3	5	379.79	3984.53	2.83	5.23	5.60
3	6	4406.71	-	12.04	38.08	26.56
3	7	49014.60	-	49.78	245.55	140.12
3	8	-	-	178.65	2348.80	913.13
3	9	-	-	703.67	15380.50	6005.73
3	10	-	-	3812.05	-	-
3	11	-	-	14246.13	-	-
3	12	-	-	56638.10	-	-
3	13	-	-	166639.09	-	-
5	2	4.18	11.85	1.67	1.46	1.90
5	3	193.05	2624.78	37.97	44.44	61.90
7	2	31.88	138.52	12.59	12.45	16.92
7	3	3499.86	105535.16	829.78	1075.83	1580.58
11	2	510.69	4550.57	188.73	211.84	316.67
13	2	1313.98	12515.75	489.52	564.15	866.52

APPENDIX B

CHOICE 1 SEQUENCES OVER $\mathbb{F}_2$

Since Choice 2, 3a, 3b and 3c do not work over $\mathbb{F}_2$ and Choice 1 can only generate very short sequences with $T_v \leq 2$, we choose the field size $p > 2$ for all analyses in this study. Proposition B.1 is concerned with the sequences generated by polynomials of Choice 1 over $\mathbb{F}_2$.

In the general description given by (1.3) of multivariate polynomial iterations,

$$F_1(\mathbf{X}) = X_1 G_1(X_2, \dots, X_m) + H_1(X_2, \dots, X_m),$$

$$\vdots$$

$$F_{m-1}(\mathbf{X}) = X_{m-1} G_{m-1}(X_m) + H_{m-1}(X_m),$$

$$F_m(\mathbf{X}) = g_m X_m + h_m,$$

Choice 1 generates the sequences by substituting

$$G_i(X_{i+1}, \dots, X_m) = X_{i+1},$$

and

$$H_i(X_{i+1}, \dots, X_m) = h_i$$

for $i = 1, ..., m-1$.

Proposition B.1. *The vector period T_v of the sequences generated by Choice 1 is less than or equal to 2 for $p = 2$.*

Proof. The vector period $T_v^{(m)}$ of the system (1.3) with m polynomials is equal to $LCM(T_{F_1}, ..., T_{F_m})$ where T_{F_i} is the vector period of each sequence generated by F_i for $i = 1, ..., m$. Let $m = 2$, then g_2 is equal to 1.

$F_1^{(0)} = X_1$ and $F_2^{(0)} = X_2$,

$F_1^{(1)} = X_1X_2 + h_1$ and $F_2^{(1)} = X_2 + h_2$,

$F_1^{(2)} = (X_1X_2 + h_1)(X_2 + h_2) + h_1 = X_1X_2^2 + h_1X_2 + h_2X_1X_2 + h_1h_2 + h_1 = X_1X_2 + h_1X_2 + h_2X_1X_2 + h_1h_2 + h_1$ and $F_2^{(2)} = X_2 + h_2 + h_2 = X_2$,

$F_1^{(3)} = (X_1X_2 + h_1X_2 + h_2X_1X_2 + h_1h_2 + h_1)X_2 + h_1 = X_1X_2^2 + h_1X_2^2 + h_2X_1X_2^2 + h_1h_2X_2 + h_1X_2 + h_1 = X_1X_2 + h_1X_2 + h_2X_1X_2 + h_1h_2X_2 + h_1X_2 + h_1$ and $F_2^{(3)} = X_2 + h_2$,

$F_1^{(4)} = (X_1X_2 + h_1X_2 + h_2X_1X_2 + h_1h_2X_2 + h_1X_2 + h_1)(X_2 + h_2) + h_1 = X_1X_2^2 + h_2X_1X_2^2 + h_1h_2X_2^2 + h_1X_2 + h_2X_1X_2 + h_2^2X_1X_2 + h_1h_2^2X_2 + h_1h_2 + h_1 = X_1X_2 + h_2X_1X_2 + h_1X_2 + h_1h_2 + h_1$ and $F_2^{(4)} = X_2 + h_2 + h_2 = X_2$

i.e. $F_1^{(2)} = F_1^{(4)}$ and $F_2^{(2)} = F_2^{(4)}$. Hence, $T_v \leq 2$.

$T_v^2 = LCM(T_{F_1}, T_{F_2}) \leq 2$ is proved. It implies that $T_{F_1} \leq 2$ and $T_{F_2} \leq 2$.

Assume $T_v \leq 2$ for any m, we will show that $T_v \leq 2$ for $m + 1$.

$T_v^{(m)} = LCM(T_{F_1}, T_{F_m})$

For $m = 2$ we proved that $T_{F_2} \leq 2$. So we only need to show that $T_{F_1} \leq 2$.

The period vector T_{F_2} of F_2 is less than or equal to 2. It implies that $F_2^{(k)} = F_2^{(k+2)}$ for $k = 0, 1, ...$

In order to compute easily, let $F_2^{(k)} = a$ when k is even and $F_2^{(k)} = b$ when k is odd.

$F_1^{(0)} = X_0$,

$F_1^{(2)} = F_1^{(0)}F_2^{(0)} + h_1$,

$F_1^{(3)} = F_1^{(1)}F_2^{(1)} + h_1$,

$F_1^{(4)} = F_1^{(2)}F_2^{(2)} + h_1$.

$F_1^{(1)} = X_1a + h_1$,

$F_1^{(2)} = (X_1a + h_1)b + h_1 = abX_1 + bh_1 + h_1$,

$F_1^{(3)} = (abX_1 + h_1b + h_1)a + h_1 = a^2bX_1 + abh_1 + ah_1 + h_1 = abX_1 + abh_1 + ah_1 + h_1$,

$F_1^{(4)} = (abX_1 + abh_1 + ah_1 + h_1)b + h_1 = ab^2X_1 + ab^2h_1 + abh_1 + bh_1 + h_1 = abX_1 + bh_1 + h_1$.

Thus, $F_1^{(2)} = F_1^{(4)}$ and $T_{F_1} \leq 2$. □

APPENDIX C

EXHAUSTIVE PERIOD ANALYSIS

Since the sequences with vector period $T_v = 1$ are equal to the initial vector $\mathbf{X} = (X_1, ..., X_m)$, their randomnesses do not depend on the method given by (1.3). As a result, they are discarded in Chapter 3, while analyzing the linear complexities. Table C.1 shows the percentage of sequences generated by five polynomial choices, whose vector period $T_v = 1$ for the examined 10 cases, corresponding to $p = 3$ with $2 \leq m \leq 5$, $p = 5, 7$ with $m = 2, 3$ and $p = 11, 13$ with $m = 2$. As can be seen from Table C.1, the sequences with $T_v = 1$ have the same percentages for Choice 2, 3a, 3b and 3c for each case. The percentages of Choice 1 sequences with vector periods equal to 1 are not the same as the other choices, but there exists not such big differences.

Additionally, the most common periods encountered in the overall space for each case and the corresponding percentages are listed in Table C.2.

Finally, the vector periods found by exhaustive search for the examined 10 cases are listed in Tables C.3-C.6. These four tables show that Choice 2 produces the largest set of vector periods, which almost always contains Choice 1, 3a, 3b and 3c sets and some extra values. On the other hand, Choice 3a, 3b and 3c have the smallest set of periods in all polynomial choices.

Table C.1: Percentage of sequences with vector period $T_v = 1$

Case	p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
1	3	2	19	11	11	11	11
2	3	3	10	4	4	4	4
3	3	4	6	1	1	1	1
4	3	5	4	0.4	0.4	0.4	0.4
5	5	2	7	4	4	4	4
6	5	3	3	1	1	1	1
7	7	2	4	2	2	2	2
8	7	3	1	0.3	0.3	0.3	0.3
9	11	2	2	1	1	1	1
10	13	2	1	1	1	1	1

Table C.2: The most common vector periods of the corresponding polynomial choices

Case	p	m	Choice 1	%	Choice 2	%	Choice 3a	%	Choice 3b	%	Choice 3c	%
1	3	2	p	41	$p-1$	33	$p(p-1)$	67	$p(p-1)$	67	$p(p-1)$	44
2	3	3	p	38	$p-1$	23	$p(p-1)$	89	$p(p-1)$	81	$p(p-1)$	74
3	3	4	p	36	$(p-1)^2$	23	$p^2(p-1)$	67	$p^2(p-1)$	44	$p(p-1)$	59
4	3	5	p	33	$(p-1)^2$	21	$p^2(p-1)$	89	$p^2(p-1)$	74	$p^2(p-1)$	49
5	5	2	$p-1$	42	$p(p-1)$	26	$p-1$	60	$p-1$	60	$p-1$	46
6	5	3	$p-1$	34	$p(p-1)$	22	$p(p-1)$	57	$p(p-1)$	62	$p(p-1)$	69
7	7	2	$p-1$	33	$p-1$	19	$p-1$	62	$p-1$	62	$p-1$	56
8	7	3	$p-1$	29	$(p-1)^2/2$	13	$p-1$	50	$p(p-1)$	47	$p(p-1)$	63
9	11	2	$p-1$	40	$(p-1)^2/2$	27	$p-1$	68	$p-1$	68	$p-1$	64
10	13	2	$p-1$	33	$(p-1)^2/3$	14	$p-1$	64	$p-1$	64	$p-1$	61

Table C.3: Vector periods obtained with two polynomials ($m = 2$)

p	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
3	1, 2, 3, 4	1, 2, 3, 4, 9	1, 2, 6	1, 2, 6	1, 2, 6
5	1, 2, 4, 5, 8, 10	1, 2, 4, 5, 8, 10, 16,20	1, 2, 4, 10, 20	1, 2, 4, 10, 20	1, 2, 4, 10, 20
7	1, 2, 3, 4, 6, 7, 9, 12, 14, 18, 21	1, 2, 3, 4, 6, 7, 9, 12, 14, 18, 21, 36, 49	1, 2, 3, 6, 14, 21, 42	1, 2, 3, 6, 14, 21, 42	1, 2, 3, 6, 14, 21, 42
11	1, 2, 4, 5, 10, 11, 20, 22, 25, 50, 55	1, 2, 4, 5, 10, 11, 20, 22, 25, 50, 55, 100, 110, 121	1, 2, 5, 10, 22, 55, 110	1, 2, 5, 10, 22, 55, 110	1, 2, 5, 10, 22, 55, 110
13	1, 2, 3, 4, 6,8, 9, 12, 13, 16, 18, 24, 26, 36, 39, 48, 72, 78	1, 2, 3, 4, 6,8, 9, 12, 13, 16, 18, 24, 26, 36, 39, 48, 52, 72, 78, 144, 156	1, 2, 3, 4, 6,12, 26, 39, 52, 78, 156	1, 2, 3, 4, 6,12, 26, 39, 52, 78, 156	1, 2, 3, 4, 6,12, 26, 39, 52, 78, 156

Table C.4: Vector periods obtained with three polynomials ($m = 3$)

p	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
3	1, 2, 3, 4, 6, 9, 12	1, 2, 3, 4, 6, 8, 9, 27	1, 2, 6	1, 2, 6	1, 2, 6
5	1, 2, 4, 5, 8, 10, 16, 20, 25, 32, 40	1, 2, 4, 5, 8, 10, 16, 20, 25, 32, 40, 64, 80, 100	1, 2, 4, 10, 20	1, 2, 4, 10, 20	1, 2, 4, 10, 20
7	1, 2, 3, 4, 6, 7, 8, 9, 12, 14, 18, 21, 24, 27, 28, 36, 42, 49, 54, 63, 72, 84	1, 2, 3, 4, 6, 7, 8, 9, 12, 14, 18, 21, 24, 27, 28, 36, 42, 49, 54, 63, 72, 84, 98, 108, 126, 147, 216, 252, 343	1, 2, 3, 6, 14, 21, 42	1, 2, 3, 6, 14, 21, 42	1, 2, 3, 6, 14, 21, 42

Table C.5: Vector periods obtained with four polynomials ($m = 4$)

p	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
3	1, 2, 3, 4, 6, 8, 9, 12	1, 2, 3, 4, 6, 8, 9, 12, 18, 27, 81	1, 2, 6, 18	1, 2, 6, 18	1, 2, 6, 18

Table C.6: Vector periods obtained with five polynomials ($m = 5$)

p	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
3	1, 2, 3, 4, 6, 9, 12	1, 2, 3, 4, 6, 8, 9, 27	1, 2, 6, 18	1, 2, 6, 18	1, 2, 6, 18

APPENDIX D

EXHAUSTIVE LINEAR COMPLEXITY ANALYSIS

In this appendix, we firstly present Tables D.1-D.12 in order to clarify the details of the LCE values given in Figures 3.4-3.7. Since the sequences with vector period $T_v = 1$ are equal to the initial vector $\mathbf{X} = (X_1, ..., X_m)$, their randomnesses do not depend on the method given by the (1.3). As a result, they are discarded in Chapter 3, while analyzing the linear complexities.

First of the three main conclusions drawn from Tables D.1-D.12 is that an increase in T_v results in serious loss of randomness for the three polynomial choices 1, 2 and 3a. The second one is that Choice 2 produces the largest set of periods, followed by Choice 1 and Choice 3 (in accordance with the results of Chapter 2, as detailed by Tables C.3-C.6). The third one is that an increase in the field size p also seems to yield some loss of randomness on the average.

Furthermore, Tables D.13 and D.14 respectively show the minimum and maximum values of the LCE of the sequences with PE ≥ 0.5, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m and h_m.

One can observe from Tables D.13 and D.14 that minimum and maximum LCE values achieved by the sequences with PE ≥ 0.5 are decreasing while the field size p and the vector size m is increasing for all polynomial choices.

Table D.1: LCE values versus vector period T_v of sequences generated by Choice 1 values for $p = 5, m = 3$

T_v	%	Minimum	Average	Maximum
2	8.9	0.50	0.93	1
4	33.7	0.25	0.79	1
5	21.4	0.40	0.83	1
8	22.2	0.42	0.84	1
10	3.2	0.40	0.73	1
16	2.8	0.44	0.48	0.50
20	1.9	0.40	0.48	0.50
25	0.8	0.40	0.40	0.40
32	1.0	0.47	0.49	0.50
40	1.5	0.26	0.36	0.40

Table D.2: LCE values versus vector period T_v of sequences generated by Choice 2 values for $p = 5, m = 3$

T_v	%	Minimum	Average	Maximum
2	2.6	0.17	0.93	1
4	10.6	0.17	0.72	1
5	2.1	0.40	0.71	1
8	12.4	0.33	0.71	1
10	2.7	0.30	0.61	1
16	15.1	0.35	0.67	1
20	22.3	0.15	0.68	1
25	0.5	0.40	0.40	0.40
32	7.7	0.35	0.52	1
40	5.4	0.37	0.74	1
64	3.1	0.35	0.47	0.50
80	8.4	0.23	0.38	0.50
100	6.4	0.38	0.39	0.40

Table D.3: LCE values versus vector period T_v of sequences generated by Choice 3a values for $p = 5, m = 3$

T_v	%	Minimum	Average	Maximum
2	2.3	0.33	0.85	1
4	32.9	0.08	0.71	1
10	7.1	0.17	0.32	0.40
20	56.9	0.10	0.18	0.20

Table D.4: LCE values versus vector period T_v of sequences generated by Choice 1 values for $p = 7, m = 3$

T_v	%	Minimum	Average	Maximum
2	3.5	0.33	0.88	1
3	10.4	0.33	0.85	1
4	1.8	0.58	0.93	1
6	29.3	0.17	0.80	1
7	12.1	0.29	0.82	1
8	0.1	0.79	0.95	1
9	6.4	0.56	0.84	1
12	12.0	0.22	0.84	1
14	2.6	0.29	0.78	1
18	8.5	0.26	0.70	1
21	5.5	0.22	0.57	1
24	0.3	0.28	0.32	0.33
27	0.2	0.56	0.65	0.67
28	0.1	0.23	0.27	0.29
36	2.8	0.26	0.56	1
42	1.8	0.29	0.32	0.33
49	0.3	0.29	0.29	0.29
54	0.2	0.28	0.32	0.33
63	0.3	0.29	0.29	0.29
72	0.4	0.32	0.33	0.33
84	0.4	0.29	0.29	0.29

Table D.5: LCE values versus vector period T_v of sequences generated by Choice 2 values for $p = 7, m = 3$

T_v	%	Minimum	Average	Maximum
2	1.2	0.17	0.88	1
3	2.2	0.33	0.81	1
4	1.1	0.50	0.92	1
6	7.2	0.17	0.73	1
7	1.3	0.19	0.75	1
8	0.3	0.83	0.96	1
9	3.2	0.44	0.74	1
12	7.2	0.22	0.76	1
14	1.2	0.19	0.63	1
18	13.3	0.22	0.63	1
21	5.2	0.14	0.66	1
24	1.4	0.28	0.67	1
27	1.8	0.48	0.64	0.67
28	0.3	0.18	0.27	0.29
36	11.1	0.19	0.57	1
42	3.6	0.07	0.40	1
49	1.1	0.27	0.49	1
54	5.4	0.24	0.47	0.67
63	5.4	0.25	0.53	0.67
72	4.6	0.21	0.40	1
84	1.1	0.21	0.26	0.29
98	0.4	0.15	0.27	0.29
108	6.6	0.31	0.31	0.31
126	6.2	0.37	0.37	0.37
147	4.5	0.27	0.40	0.67
216	1.1	0.22	0.32	0.33
252	1.2	0.16	0.21	0.29
343	0.8	0.28	0.28	0.29

Table D.6: LCE values versus vector period T_v of sequences generated by Choice 3a values for $p = 7, m = 3$

T_v	%	Minimum	Average	Maximum
2	0.9	0.20	0.80	1
3	3.7	0.30	0.77	1
6	50.1	0.20	0.54	0.70
14	2.0	0.10	0.19	0.20
21	7.4	0.10	0.16	0.20
42	35.7	0	0.10	0.10

Table D.7: LCE values versus vector period T_v of sequences generated by Choice 1 values for $p = 11, m = 2$

T_v	%	Minimum	Average	Maximum
2	3.1	0.50	0.96	1
4	0.8	0.75	0.98	1
5	21.3	0.20	0.80	1
10	40.3	0.10	0.63	1
11	9.8	0.18	0.87	1
20	6.0	0.15	0.58	1
22	0.6	0.18	0.18	0.18
25	7.5	0.28	0.38	0.40
50	7.5	0.14	0.18	0.20
55	1.5	0.18	0.18	0.18

Table D.8: LCE values versus vector period T_v of sequences generated by Choice 2 values for $p = 11, m = 2$

T_v	%	Minimum	Average	Maximum
2	1.8	0.50	0.96	1
4	0.9	0.88	0.98	1
5	6.6	0.20	0.63	1
10	10.5	0.10	0.48	1
11	1.6	0.18	0.57	1
20	3.6	0.18	0.20	0.20
22	0.7	0.14	0.18	0.18
25	12.0	0.28	0.37	0.40
50	27.0	0.12	0.22	0.40
55	9.3	0.15	0.32	0.40
100	18.0	0.12	0.18	0.20
110	5.4	0.15	0.18	0.18
121	1.7	0.18	0.18	0.18

Table D.9: LCE values versus vector period T_v of sequences generated by Choice 3a values for $p = 11, m = 2$

T_v	%	Minimum	Average	Maximum
2	1.7	0.50	0.94	1
5	19.2	0.10	0.49	0.60
10	68.2	0.10	0.28	0.30
22	1.1	0.10	0.10	0.10
55	4.5	0	0.09	0.10
110	4.5	0	0	0

Table D.10: LCE values versus vector period T_v of sequences generated by Choice 1 values for $p = 13, m = 2$

T_v	%	Minimum	Average	Maximum
2	2.2	0.50	0.93	1
3	5.6	0.33	0.87	1
4	7.1	0.25	0.86	1
6	10.0	0.17	0.78	1
8	1.5	0.38	0.60	1
9	2.5	0.56	0.64	0.67
12	32.8	0.08	0.61	1
13	8.2	0.15	0.88	1
16	2.9	0.38	0.48	0.50
18	2.5	0.28	0.32	0.33
24	5.1	0.13	0.53	1
26	0.5	0.15	0.15	0.15
36	4.9	0.14	0.16	0.17
39	0.3	0.15	0.15	0.15
48	5.8	0.13	0.16	0.17
72	6.6	0.11	0.15	0.17
78	0.5	0.15	0.15	0.15

Table D.11: LCE values versus vector period T_v of sequences generated by Choice 2 values for $p = 13, m = 2$

T_v	%	Minimum	Average	Maximum
2	1.2	0.25	0.94	1
3	2.4	0.33	0.79	1
4	3.1	0.25	0.73	1
6	4.9	0.17	0.70	1
8	2.4	0.38	0.67	1
9	2.5	0.44	0.64	0.67
12	12.4	0.08	0.55	1
13	1.1	0.15	0.56	1
16	2.7	0.38	0.48	0.50
18	5.2	0.22	0.45	0.67
24	11.3	0.13	0.55	1
26	0.5	0.12	0.15	0.15
36	11.5	0.11	0.36	0.67
39	1.2	0.12	0.14	0.15
48	14.1	0.13	0.35	0.50
52	3.3	0.10	0.38	0.50
72	9.8	0.10	0.23	0.33
78	0.9	0.13	0.14	0.14
144	4.4	0.10	0.16	0.17
156	4.7	0.09	0.16	0.17

Table D.12: LCE values versus vector period T_v of sequences generated by Choice 3a values for $p = 13, m = 2$

T_v	%	Minimum	Average	Maximum
2	1.2	0.30	0.93	1
3	3.9	0.20	0.76	1
4	6.7	0.30	0.67	0.80
6	15.1	0.10	0.43	0.50
12	64.1	0.10	0.27	0.30
26	0.8	0.10	0.10	0.10
39	1.5	0.10	0.10	0.10
52	1.5	0	0.09	0.10
78	1.5	0	0	0
156	3.1	0	0	0

Table D.13: Minimum linear complexity efficiency of the sequences with period efficiency PE ≥ 0.5, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m and h_m

Case	p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
1	3	2	-	0.67	0.33	0.33	0.42
2	3	3	-	0.67	-	-	-
3	3	4	-	0.67	-	-	-
4	3	5	-	0.67	-	-	-
5	5	2	-	0.30	0.10	0.10	0.15
6	5	3	-	0.23	-	-	-
7	7	2	-	0.22	0	0.05	0.07
8	7	3	-	0.16	-	-	-
9	11	2	0.18	0.18	0	0.02	0.03
10	13	2	0.15	0.09	0	0.01	0.02

Table D.14: Maximum linear complexity efficiency of the sequences with period efficiency PE ≥ 0.5, found over all possible values of $\mathbf{X}$, $\mathbf{G}$, $\mathbf{H}$, a_i, g_m and h_m

Case	p	m	Choice 1	Choice 2	Choice 3a	Choice 3b	Choice 3c
1	3	2	-	0.67	0.42	0.42	0.42
2	3	3	-	0.67	-	-	-
3	3	4	-	0.67	-	-	-
4	3	5	-	0.67	-	-	-
5	5	2	-	0.50	0.20	0.15	0.20
6	5	3	-	0.50	-	-	-
7	7	2	-	0.33	0.10	0.07	0.10
8	7	3	-	0.29	-	-	-
9	11	2	0.18	0.18	0.10	0.05	0.04
10	13	2	0.15	0.17	0	0.06	0.03

APPENDIX E

LINEAR COMPLEXITY ANALYSIS OF THE LARGEST PERIOD SEQUENCES

In this appendix, we present some tables on minimum, average and maximum linear complexity efficiencies attained within the set of the largest period sequences of the corresponding polynomial choices (1, 2, 3a, 3b and 3c)

i) versus the field size $3 \leq p \leq 11$ and 13 with the number of polynomials, $m = 2$ in Tables E.1-E.5;

ii) versus the vector size m up to 7, 5, 13, 9 and 9 for Choices 1, 2, 3a, 3b and 3c where $p = 3$ in Tables E.6-E.10.

As can be seen from Tables E.1-E.5, increasing the field size p from 3 to 13 causes all five polynomial choices to generate sequences with lower and still lower LCE values. Despite the fact that Choice 1 sequences seem slightly better than Choice 2 sequences, which are much better than Choice 3 sequences, none of the LCE values in these tables exceed 0.95, except the maximum LCE of Choice 1 for $p = 3$.

Table E.6 shows that the LCE values of the largest period sequences generated by Choice 1 in $\mathbb{F}_3$ are at most 0.67 for odd m, but they can achieve the highest point (LCE ≥ 0.95) when m is even.

The other observation from Table E.7 is that for the field size $p = 3$, LCE values of the largest period sequences produced by Choice 2 are 0.67 independently of m.

Table E.8 indicates that LCE values of the sequences having largest period generated by Choice 3a in $\mathbb{F}_3$ can be higher when m is a power of $p = 3$.

On the other hand, Tables E.9 and E.10 demonstrate that the LCE values of the largest period sequences of Choices 3b and 3c for the field size $p = 3$ are quite similar, and they seem better than that of Choice 3a.

Table E.1: LCE values obtained at the largest T_v of Choice 1 at $m = 2$

p	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
3	4	0.44	7.4	0.75	0.92	1
5	10	0.40	1.6	0.40	0.40	0.40
7	21	0.43	1.7	0.29	0.29	0.29
11	55	0.45	1.5	0.18	0.18	0.18
13	78	0.46	0.5	0.15	0.15	0.15

Table E.2: LCE values obtained at the largest T_v of Choice 2 at $m = 2$

p	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
3	9	1	22.2	0.67	0.67	0.67
5	20	0.80	25.6	0.30	0.42	0.50
7	49	1	2.7	0.29	0.29	0.29
11	121	1	1.7	0.18	0.18	0.18
13	156	0.92	4.7	0.09	0.16	0.17

Table E.3: LCE values obtained at the largest T_v of Choice 3a at $m = 2$

p	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
3	6	0.67	66.7	0.33	0.39	0.42
5	20	0.80	17.8	0.10	0.14	0.20
7	42	0.86	6.9	0	0.09	0.10
11	110	0.91	4.5	0	0	0
13	156	0.92	3.1	0	0	0

Table E.4: LCE values obtained at the largest T_v of Choice 3b at $m = 2$

p	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
3	6	0.67	66.7	0.33	0.39	0.42
5	20	0.80	17.8	0.10	0.13	0.15
7	42	0.86	6.9	0.05	0.07	0.07
11	110	0.91	4.5	0.02	0.03	0.03
13	156	0.92	3.1	0.01	0.02	0.04

Table E.5: LCE values obtained at the largest T_v of Choice 3c at $m = 2$

p	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
3	6	0.67	44.4	0.42	0.42	0.42
5	20	0.80	32	0.15	0.18	0.20
7	42	0.86	12.7	0.07	0.09	0.10
11	110	0.91	8.6	0.03	0.04	0.04
13	156	0.92	5.9	0.02	0.02	0.03

Table E.6: LCE values obtained at the largest T_v of Choice 1 versus m for $p = 3$

m	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
2	4	0.44	7.4	0.75	0.92	1
3	12	0.44	1.7	0.67	0.67	0.67
4	12	0.15	2.9	0.67	0.79	1
5	36	0.15	0.3	0.61	0.63	0.67
6	36	0.05	0.4	0.61	0.71	0.97
7	108	0.05	0.0	0.65	0.66	0.67

Table E.7: LCE values obtained at the largest T_v of Choice 2 versus m for $p = 3$

m	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
2	9	1	22.2	0.67	0.67	0.67
3	27	1	14.8	0.67	0.67	0.67
4	81	1	9.9	0.67	0.67	0.67
5	243	1	6.6	0.67	0.67	0.67

Table E.8: LCE values obtained at the largest T_v of Choice 3a versus m for $p = 3$

m	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
2	6	0.67	66.7	0.33	0.39	0.42
3	6	0.22	88.9	0.33	0.57	0.67
4	18	0.22	66.7	0.22	0.23	0.24
5	18	0.07	88.9	0.22	0.30	0.33
6	18	0.02	96.3	0.22	0.32	0.34
7	18	0.01	98.8	0.22	0.40	0.44
8	18	3.E-03	99.6	0.22	0.42	0.45
9	18	9.E-04	99.9	0.22	0.51	0.56
10	54	9.E-04	66.7	0.19	0.19	0.19
11	54	3.E-04	88.9	0.19	0.21	0.22
12	54	1.E-04	96.3	0.19	0.22	0.22
13	54	3.E-05	98.8	0.19	0.24	0.26

Table E.9: LCE values obtained at the largest T_v of Choice 3b versus m for $p = 3$

m	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
2	6	0.67	66.7	0.33	0.39	0.42
3	6	0.22	81.5	0.50	0.73	0.83
4	18	0.22	44.4	0.39	0.39	0.39
5	18	0.07	74.1	0.37	0.59	0.72
6	54	0.07	44.4	0.35	0.35	0.35
7	54	0.02	64.2	0.35	0.58	0.69
8	162	0.02	44.4	0.34	0.34	0.34
9	162	0.01	64.2	0.34	0.57	0.67

Table E.10: LCE values obtained at the largest T_v of Choice 3c versus m for $p = 3$

m	at Max. T_v	Corresponding PE	%	Min. LCE	Ave. LCE	Max. LCE
2	6	0.67	44.4	0.42	0.42	0.42
3	6	0.22	74.1	0.44	0.79	1
4	18	0.22	29.6	0.38	0.38	0.38
5	18	0.07	49.4	0.36	0.62	0.78
6	54	0.07	14.8	0.35	0.35	0.35
7	54	0.02	24.7	0.35	0.56	0.70
8	162	0.02	9.9	0.33	0.33	0.34
9	162	0.01	19.8	0.33	0.50	0.68

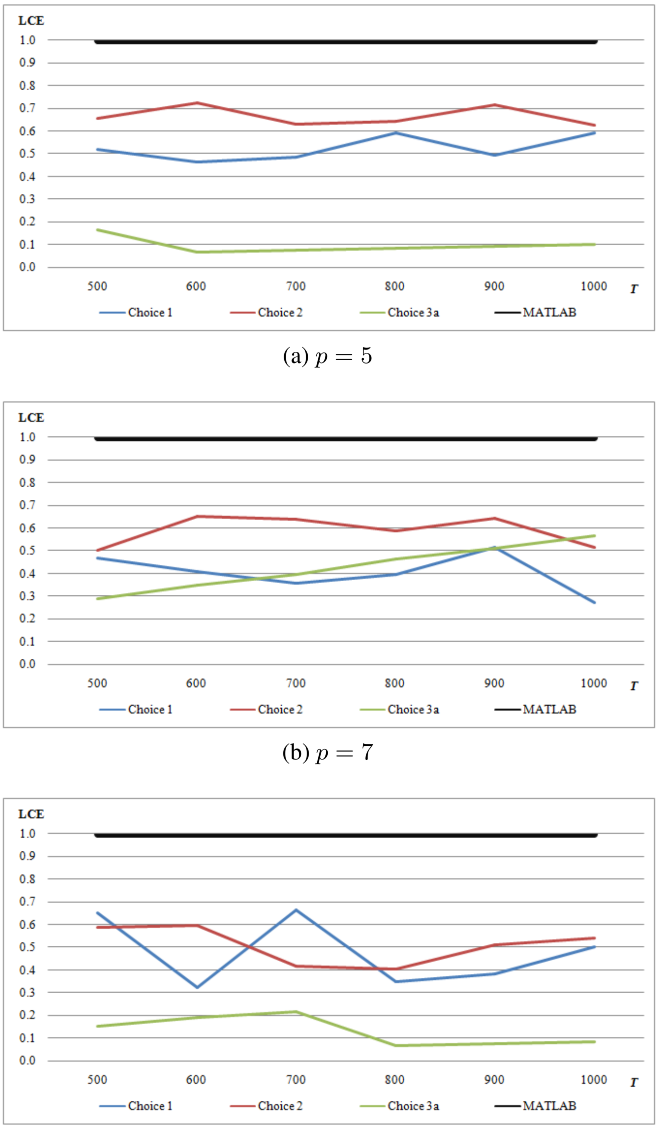
APPENDIX F

LINEAR COMPLEXITIES OF SEQUENCES GENERATED WITH RANDOM INITIAL VALUES

In this appendix, we present some curves and graphs on the linear complexity efficiency, LCE, as defined in this work by "linear complexity L divided by the period T of the sequence". Figure F.1 is sketched for $5 \leq p \leq 11$, in addition to Figure 5.1 that shows the LCE values for $p = 3$ and 13. Similarly, Figure F.2 is sketched for $5 \leq p \leq 11$, in addition to Figure 5.2 that shows the percentage of sequences with LCE > 0.95 for $p = 3$ and 13.

Figure F.1 indicates that the average LCE values of the reference sequences generated by MATLAB look perfectly random, with average LCE almost equal to 1. On the other hand, Ostafe's sequences, Choice 1, 2 and 3a, have lower LCE values for all T 's in the given set for $5 \leq p \leq 11$, since their average LCE values do not reach 0.8.

Figure F.2 shows that for field sizes $5 \leq p \leq 11$, Choice 1 and 2 generate highly random sequences with very low percentages (not exceeding 50%).

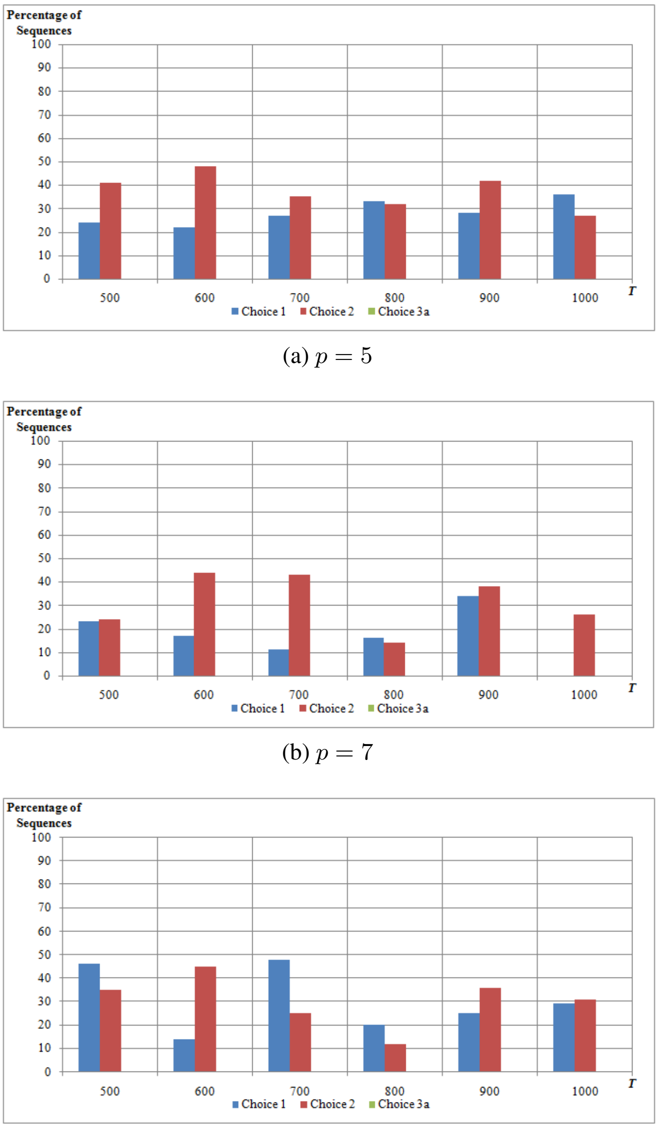


(a) $p = 5$

(b) $p = 7$

(c) $p = 11$

Figure F.1: Randomness comparison of the three choices with the random sequences produced by MATLAB's **randi(.)** where (a) $p = 5$, (b) $p = 7$, (c) $p = 11$, in terms of the LCE as defined in this work by "linear complexity L divided by the period T of the sequence" (For each T, average LCE is computed over 100 sequences.)



(a) $p = 5$

(b) $p = 7$

(c) $p = 11$

Figure F.2: Percentage of sequences with LCE ≥ 0.95 for (a) $p = 5$, (b) $p = 7$, (c) $p = 11$ using three choices

APPENDIX G

PERCENTAGES OF HIGH-LCE SEQUENCES GENERATED WITH RANDOM INITIAL VALUES AND CORRESPONDING PE'S

In Section 5.5, we generate 100 sequences with random initialization for $3 \leq p \leq 43$ and $2 \leq m \leq 4$. We then compute the average LCE's of the corresponding sequences, whose vector periods $T_v \geq p-1$. For each set of 100 sequences, this appendix presents the percentages of those with LCE ≥ 0.95 and corresponding period efficiencies in Tables G.1, G.2 and G.3 for $m = 2, 3$ and 4 respectively.

Average LCE values of the sequences generated by all five choices decreases with an increase of field size p. For the field sizes $3 \leq p \leq 43$ with vector sizes $m = 2, 3$ and the field sizes $3 \leq p \leq 19$ with vector sizes $m = 4$, Choice 1 and 2 can generate sequences with high LCE (≥ 0.95) but the corresponding PE is less than or equal to 0.44 for $m = 2$ and less than or equal to 0.33 for $m = 3, 4$. On the other hand, Choice 3a and 3b can not produce high-LCE sequences for $m = 2$; however, Choice 3c can generate such sequences for only small field sizes ($p = 3$ and 5) but the corresponding PE is less than or equal to 0.22. In addition, for vector size $m = 3$, Choice 3a, 3b and 3c can produce high-LCE sequences for field size p not exceeding 3, 7 and 7 and period efficiency PE not exceeding 0.07, 0.07 and 0.22, respectively. Similarly, for $m = 4$, Choice 3a, 3b and 3c generate high-LCE sequences mainly for small field sizes but with very low period efficiencies.

Table G.1: Percentages of high-LCE sequences (LCE ≥ 0.95) with $T_v \geq p-1$ and corresponding maximum PE's generated by random initialization at $m=2$

p	Choice 1	Maximum PE	Choice 2	Maximum PE	Choice 3a	Maximum PE
3	69	0.44	5	0.04	0	-
5	52	0.32	35	0.44	0	-
7	31	0.24	4	0.16	0	-
11	29	0.17	10	0.24	0	-
13	22	0.14	3	0.09	0	-
17	23	0.11	10	0.14	0	-
19	16	0.05	3	0.11	0	-
23	16	0.08	0	-	0	-
29	14	0.07	13	0.08	0	-
31	12	0.03	2	0.03	0	-
37	7	0.05	1	0.03	0	-
41	9	0.05	5	0.05	0	-
43	12	0.02	1	0.02	0	-

Table G.2: Percentages of high-LCE sequences (LCE ≥ 0.95) with $T_v \geq p-1$ and corresponding maximum PE's generated random initialization at $m=3$

p	Choice 1	Maximum PE	Choice 2	Maximum PE	Choice 3a	Maximum PE
3	51	0.22	32	0.33	5	0.07
5	47	0.08	32	0.32	3	0.03
7	34	0.06	10	0.07	0	-
11	53	0.04	11	8.E-02	0	-
13	50	0.04	11	7.E-02	0	-
17	51	0.03	14	1.E-01	0	-
19	61	0.02	8	3.E-02	0	-
23	53	0.02	11	4.E-02	0	-
29	57	0.02	9	3.E-02	0	-
31	64	0.02	7	3.E-02	0	-
37	63	0.01	4	3.E-02	0	-
41	68	0.01	4	2.E-02	0	-
43	67	0.01	2	2.E-02	0	-

Table G.3: Percentages of high-LCE sequences (LCE ≥ 0.95) with $T_v \geq p-1$ and corresponding maximum PE's generated by random initialization at $m = 4$

p	Choice 1	Maximum PE	Choice 2	Maximum PE	Choice 3a	Maximum PE
3	30	0.10	28	0.33	0	-
5	42	0.06	22	0.16	1	0.01
7	50	0.03	30	0.18	0	-
11	78	0.01	16	0.14	0	-
13	74	0.01	18	0.06	0	-
17	75	0.01	4	0.05	0	-
19	71	0.00	11	0.05	0	-

APPENDIX H

LCE VALUES OF THE SEQUENCES GENERATED BY CHOICE 3a WITH RANDOM INITIAL VALUES FOR $T = 1000$

Table 5.2 in Section 5.4 shows the LCE values of the sequences with period $T = 1000$ generated by Choice 3a for $p = 3$. In this appendix, we extend these results to $5 \leq p \leq 31$.

Table H.1-H.9 show that the LCE values of the sequences generated by Choice 3a vary in a narrow range. These tables show the slight increase in the average LCE values corresponding to the slight increase in the number of polynomials is not a sufficient indicator to claim some correlation between m and LCE parameters.

Table H.1: LCE values of the randomly initiated sequences generated by Choice 3a for $p = 5, T = 1000$

m	Minimum	Average	Maximum	Count
10	0.08	0.10	0.11	100

Table H.2: LCE values of the randomly initiated sequences generated by Choice 3a for $p = 7, T = 1000$

m	Minimum	Average	Maximum	Count
23	0.48	0.54	0.57	30
24	0.50	0.56	0.60	39
25	0.55	0.59	0.62	31

Table H.3: LCE values of the randomly initiated sequences generated by Choice 3a for $p = 11$, $T = 1000$

m	Minimum	Average	Maximum	Count
9	0.06	0.08	0.09	100

Table H.4: LCE values of the randomly initiated sequences generated by Choice 3a for $p = 13$, $T = 1000$

m	Minimum	Average	Maximum	Count
13	0.12	0.16	0.18	100

Table H.5: LCE values of the randomly initiated sequences generated by Choice 3a for $p = 17$, $T = 1000$

m	Minimum	Average	Maximum	Count
7	0.04	0.05	0.06	100

Table H.6: LCE values of the randomly initiated sequences generated by Choice 3a for $p = 19$, $T = 1000$

m	Minimum	Average	Maximum	Count
3	0.01	0.01	10.01	88
6	0.03	0.04	0.04	12

Table H.7: LCE values of the randomly initiated sequences generated by Choice 3a for $p = 23$, $T = 1000$

m	Minimum	Average	Maximum	Count
2	0.00	0.01	0.01	47
4	0.01	0.02	0.02	53

Table H.8: LCE values of the randomly initiated sequences generated by Choice 3a for $p = 29$, $T = 1000$

m	Minimum	Average	Maximum	Count
5	0.02	0.03	0.03	100

Table H.9: LCE values of the randomly initiated sequences generated by Choice 3a for $p = 31$, $T = 1000$

m	Minimum	Average	Maximum	Count
2	0.00	0.01	0.01	56
3	0.01	0.01	0.01	43
5	0.03	0.03	0.03	1

APPENDIX I

MATLAB IMPLEMENTATIONS OF ALGORITHMS IN THIS WORK

I.1 MATLAB Implementation of Choice 1

```
function [Tv,U,UPeriodic,URemain]=Choice1(p,m,gm,hm,H,X)
Tv=1;
U=X;
while(Tv =p m)
    F=ones(1,m);
    for r=1:m 1
        F(r)=mod(X(r) X(r+1)+H(r),p);
    end
    F(m)=mod(gm X(m)+hm,p);
    for i=1:m:(Tv 1 ) m+1
        if(U(i:i+m 1)==F)
            Tv=(Tv m+1 i)/m;
            UPeriodic=U(i:i+(Tv m) 1);
            URemain=U(1:i 1);
            return;
        end
    end
    U=[U F];
    X=F;
    Tv=Tv+1;
end
```

I.2 MATLAB Implementation of Choice 2

```
function [Tv,U,UPeriodic,URemain]=Choice2(m,p,gm,hm,H,A,X)
Tv=1;
U=X;
while(Tv =p m)
    F=ones(1,m);
    for r=1:m 1
        F(r)=mod(X(r) (X(r+1) 2 A(r))+H(r),p);
    end
    F(m)=mod(gm X(m)+hm,p);
    for i=1:m:(Tv 1 ) m+1
        if(U(i:i+m 1)==F)
            Tv=(Tv m+1 i)/m;
            UPeriodic=U(i:i+(Tv m) 1);
            URemain=U(1:i 1);
            return;
        end
    end
    U=[U F];
    X=F;
    Tv=Tv+1;
end
```

I.3 MATLAB Implementation of Choice 3a

```
function [Tv,U,UPeriodic,URemain]=Choice3(m,p,gm,hm,G,X)
Tv=1;
U=X;
while(Tv =p m)
    F=ones(1,m);
    for r=1:m 1
        F(r)=mod(X(r) G(r)+X(r+1),p);
    end
    F(m)=mod(gm X(m)+hm,p);
    for i=1:m:(Tv 1 ) m+1
        if(U(i:i+m 1)==F)
            Tv=(Tv m+1 i)/m;
            UPeriodic=U(i:i+(Tv m) 1);
            URemain=U(1:i 1);
            return;
        end
    end
    U=[U F];
    X=F;
    Tv=Tv+1;
end
```

I.4 MATLAB Implementation of Choice 3b

```
function [Tv,U,UPeriodic,URemain]=Choice3b(m,p,gm,hm,G,X)
Tv=1;
U=X;
while(Tv =p m)
    F=ones(1,m);
    for r=1:m 1
        F(r)=mod(X(r) G(r)+prod(X(r+1:m)),p);
    end
    F(m)=mod(gm X(m)+hm,p);
    for i=1:m:(Tv 1) m+1
        if(U(i:i+m 1)==F)
            Tv=(Tv m+1 i)/m;
            UPeriodic=U(i:i+(Tv m) 1);
            URemain=U(1:i 1);
            return;
        end
    end
    U=[U F];
    X=F;
    Tv=Tv+1;
end
```

I.5 MATLAB Implementation of Choice 3c

```
function [Tv,U,UPeriodic,URemain]=Choice3c(m,p,gm,hm,G,X)
Tv=1;
U=X;
while(Tv =p m)
    F=ones(1,m);
    for r=1:m 1
        F(r)=mod(X(r) G(r)+(X(r+1)) 2,p);
    end
    F(m)=mod(gm X(m)+hm,p);
    for i=1:m:(Tv 1) m+1
        if(U(i:i+m 1)==F)
            Tv=(Tv m+1 i)/m;
            UPeriodic=U(i:i+(Tv m) 1);
            URemain=U(1:i 1);
            return;
        end
    end
    U=[U F];
    X=F;
    Tv=Tv+1;
end
```

I.6 MATLAB Implementation of Choice 4

```
function Tv=Choice4(m,p,hm,X)
Tv=1;
U=X;
while(Tv =p m)
    F=zeros(1,m);
    for r=1:m 1
        product=1;
        for c=r+1:m
            product=product p1thPowerMod(X(c));
        end
        F(r)=mod(X(r)+product,p);
    end
    F(m)=mod(X(m)+hm,p);
    for i=1:m:(Tv 1 ) m+1
        if(U(i:i+m 1)==F)
            Tv=(Tv m+1 i)/m;
            return;
        end
    end
    U=[U F];
    X=F;
    Tv=Tv+1;
end
```

I.7 MATLAB Implementation of the Berlekamp-Massey Algorithm

```
function [L,LP,c]=BM(s,p)
n=length(s);
L=0;
LP=zeros(1,n);
c=zeros(1,n);
c(1)=1;
c1=zeros(1,n);
c1(1)=1;
t=zeros(1,n);
e=1;
d=0;
d1=1;
for i=0:n 1
    d=0;
    for j=0:L
        d=mod(d+c(j+1) s(i j+1),p);
    end
    if(d==0)
        e=e+1;
    elseif (2 L i)
        temp=zeros(1,n);
        temp(e+1)=gfdiv(d,d1,p);
        c=gfsub(c,gfconv(temp,c1,p),p);
        e=e+1;
    else
        L=i+1 L;
        t=c;
        temp=zeros(1,n);
        temp(e+1)=gfdiv(d,d1,p);
        c=gfsub(c,gfconv(temp,c1,p),p);
        c1=t;
        d1=d;
        e=1;
    end
    LP(i+1)=L;
end
c=c(1:L+1);
```

APPENDIX J

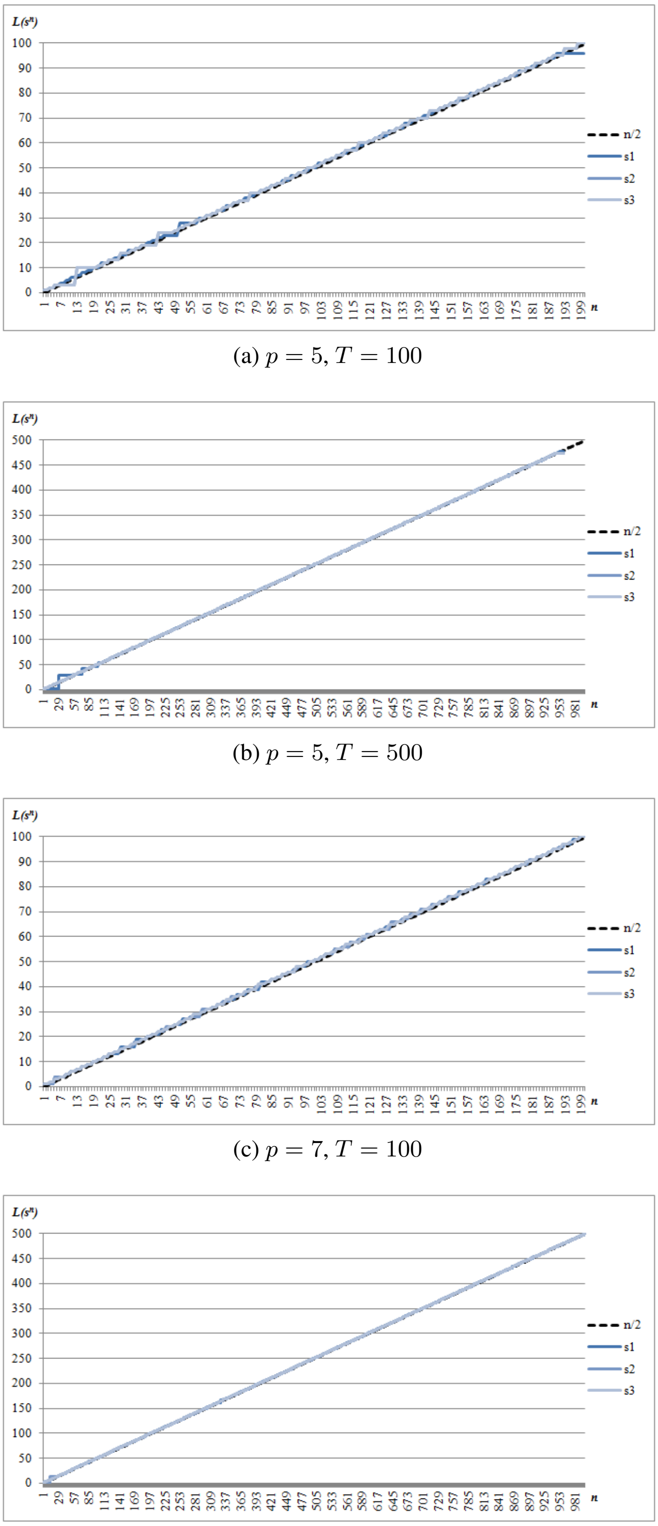
LINEAR COMPLEXITY PROFILES

As it is mentioned before in Section 1.3, the linear complexity profile of a random sequence should increase approximately as the $n/2$ line. It means that the linear complexity profile of a random sequence should be close to $n/2$ line for $n = 1, 2, \ldots, 2T$, and achieve the period T after $2T$ terms. In order to investigate whether the sequences generated by all five polynomial choices, their linear complexities are calculated after each term of the sequences.

In this appendix we draw the linear complexity profiles of some example sequences, by assigning fix period T (within the range $T \pm 0.05T$) and field size p with random values to the remaining parameters (i.e., the number of polynomials m, $\mathbf{X} = (X_1, ..., X_m)$, $\mathbf{G} = (G_1, ..., G_{m-1})$, $\mathbf{H} = (H_1, ..., H_{m-1})$, a_i, g_m and h_m) of the related polynomial choice.

Figure J.1 shows the linear complexity profile of high LCE sequences of length $T = \{100 \pm 0.05, 500 \pm 0.05\}$ generated by Choice 1 for field sizes $p = 5$ and $p = 7$. It is observed that the sequences with high LCE produced by using the polynomials in Choice 1 have the linear complexity profiles, which are close to $n/2$ line.

On the other hand, Figure J.2 show that the linear complexity profiles of some sequences with LCE < 0.95 generated by Choice 1 for field size $p = 5$ and 7 as an example. It can be observed that the linear complexity profile of low LCE sequences are not close to $n/2$ line for $n = 1, 2, \ldots, 2T$, and achieve the period T after $2T$ terms.



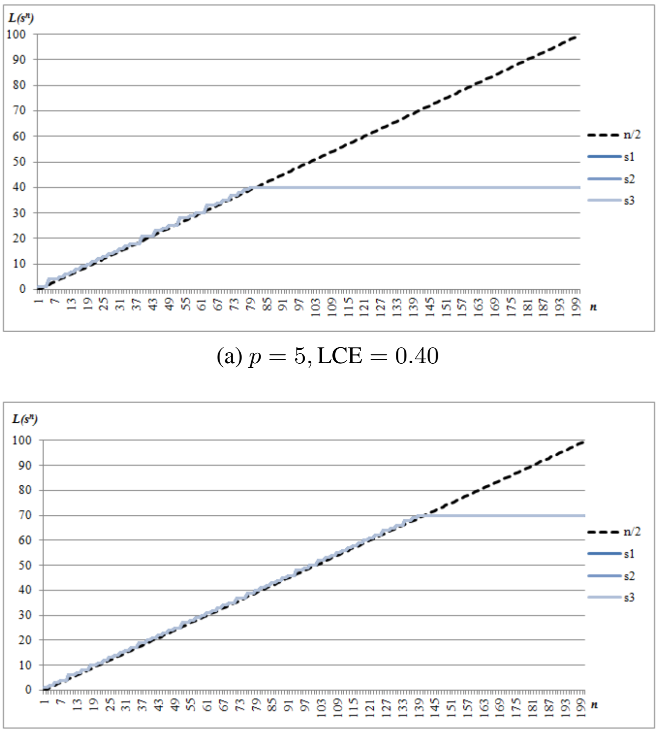
(a) $p = 5, T = 100$

(b) $p = 5, T = 500$

(c) $p = 7, T = 100$

(d) $p = 7, T = 500$

Figure J.1: Linear complexity profile of three sequences generated by Choice 1 with high (LCE ≥ 0.95) where the field size p and the scalar period T are given as (a) $p = 5, T = 100$, (b) $p = 5, T = 100$, (c) $p = 7, T = 500$, (d) $p = 7, T = 500$



(a) $p = 5, \text{LCE} = 0.40$

(b) $p = 7, \text{LCE} = 0.67$

Figure J.2: Linear complexity profile of three sequences with LCE < 0.7 generated by Choice 1 of the scalar period $T = 100$ where the field size p is given as (a) $p = 5$, (b) $p = 7$

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Gürkan Balıkçıoğlu, Pınar
Nationality: Turkish
Date and Place of Birth: 29 September 1982, Ankara
Marital Status: Married

EDUCATION

Degree	**Institution**	**Year of Graduation**
B.S.	Hacettepe Univeristy, Department of Statistics	2004
High School	Ankara Aydınlıkevler Anatolian High School	2000

PROFESSIONAL EXPERIENCE

Year	**Place**	**Enrollment**
2006-2007	METU, Graduate School of Informatics	Assistant Student
2007-Present	Research and Development Institution, Ankara, TURKEY	Specialist

PUBLICATIONS

International Conference Publications

- P.G. Balıkçıoğlu, M.D. Yücel. *Period Analysis Of Pseudorandom Vector Sequences With Dynamical Polynomial Systems.* The 6-th International Conference on Information Security and Cryptology, Ankara, Turkey, 2013.
- P.G. Balıkçıoğlu, M.D. Yücel. *Randomness Properties of Some Vector Sequences Generated by Multivariate Polynomial Iterations.* The 12-th International Conference on Finite Fields and Their Applications (Book of Abstracts), Saratoga, NY, USA, 2015.