On the Number of Binary Sequences with Good Linear Complexity Profiles

Markus Schneider and Oliver Stutzke
Fachgebiet Kommunikationssysteme
FernUniversität Hagen, 58084 Hagen
e-mail: {mark.schneider | oliver.stutzke}@fernuni-hagen.de

***Abstract* — The linear complexity profile of pseudorandom sequences to be used in stream cipher systems provides a criterion concerning their element's unpredictability which is necessary for secure encryption. In this paper, we present new results concerning the number of finite sequences with good linear complexity profiles and determine the probability for an arbitrary chosen finite sequence with given length to have good linear complexity profile.**

I. Introduction

Technically relevant stream cipher systems use finite binary pseudorandom sequences $s^n = s_0, s_1, \ldots, s_{n-1}$, with $s_m \in GF(2)$ for $0 \leq m < n$ and $m, n \in \mathbb{Z}$. These sequences are produced by a pseudorandom generator and the generation process is controlled by a cryptographic key. For encryption, the pseudorandom sequence is added componentwisely to the binary elements of the plaintext message. In order to decrypt the encrypted message, the same pseudorandom sequence has to be substracted componentwisely from the obtained ciphertext. In the following, we restrict our considerations to binary sequences.

For security reasons, the pseudorandom generator should produce the sequence in such a way that an attacker should not be able to predict any unknown sequence element with probability better than $P(s_m = 0) = P(s_m = 1) = 0.5$ even if he knows the generation principle or any other sequence elements. Therefore, as a necessary condition for unpredictability, a sequence should have high complexity. A sequence complexity measure indicates the difficulty in predicting the sequence. In this context, the linear complexity $L(s^n)$ of a finite sequence s^n is very important. The linear complexity is defined as the length of the shortest linear feedback shift-register (LFSR) that can produce s^n. In general, $0 \leq L(s^n) \leq n$. The generating LFSR and $L(s^n)$ can be found by the Berlekamp-Massey algorithm if the attacker knows $2 \cdot L(s^n)$ sequence elements [1]. Thus, for the avoidance of the sequence reconstruction by means of some known sequence elements and the application of the Berlekamp-Massey algorithm, the linear complexity should be high-valued.

Rueppel showed in his work [3] that a high-valued linear complexity is not sufficient for unpredictability and introduced the linear complexity profile of a sequence s^n as $L(s^1), L(s^2), \ldots, L(s^n)$. He also stated that the linear complexity $L(s^m)$ should follow the $\frac{m}{2}$-line in a close manner for $1 \leq m \leq n$.

Similarly as Niederreiter did in [2], we define a sequence s^n to have a good linear complexity profile, if for a real number $K \geq 1$ and $2 \leq m \leq n$ the linear complexity is bounded by

$$\lceil \frac{m}{2} \rceil - K \cdot \log_2 m \leq L(s^m) \leq \lceil \frac{m}{2} \rceil + K \cdot \log_2 m. \quad (1)$$

Note, that in this definition there is no restriction concerning the linear complexity $L(s^1)$ of the start segment consisting only of the first sequence element. We exclusively consider values $K \geq 1$. If $0 < K < 1$, we would obtain sequences with start segments that are partially predictable by a property found by Wang [5].

In general, the number of sequences that fulfill certain security requirements is of high interest. In this context, the number of finite sequences with length n and $L(s^n) = l$ was given in [3]. Furthermore, the number of finite sequences with length n and never exceeding a distance $\delta > 0$ from $\lceil \frac{m}{2} \rceil$ for $1 \leq m \leq n$ can be found in [4]. In this paper, we give the number of finite sequences with length n and good linear complexity profile as it was defined in expression (1).

In section II, we introduce some necessary fundamentals and properties in the area of linear complexities. In section III, we give the new result concerning the number of binary sequences with good linear complexity profile. Furthermore, we consider the probability to find a sequence with good linear complexity profile in the set of all finite binary sequences of length n.

II. Relevant Properties of Linear Complexity

The properties of the linear complexity follow mainly from [1]. Let $1 \leq m \leq n$. If the shortest LFSR which is able to produce s^{m-1} also generates the next element s_{m-1} then $L(s^m) = L(s^{m-1})$. If the LFSR is not able to produce s_{m-1} then $L(s^m) = \max(L(s^{m-1}), m - L(s^{m-1}))$. In this case, we have

$$L(s^m) = \begin{cases} L(s^{m-1}), \text{ if } L(s^{m-1}) \geq \frac{m}{2} \\ m - L(s^{m-1}), \text{ if } L(s^{m-1}) < \frac{m}{2}. \end{cases} \quad (2)$$

This property of linear complexity is indicated in figure 1. Each point in this figure represents all sequences of given length with specific linear complexity. The numbers beside the arrows show the numbers of possibilities to obtain a sequence with m elements and linear complexity $L(s^m)$ from a sequence with $m-1$ elements and linear complexity $L(s^{m-1})$.

The number of binary sequences $N_n(l)$ with n elements and linear complexity $L(s^n) = l$ was given in [3] as

$$N_n(l) = \begin{cases} 2^{\min(2n-2l,2l-1)}, \text{ if } 0 < l \leq n \\ 1, \text{ if } l = 0 \leq n. \end{cases} \quad (3)$$

Analyzing expression (3), one finds that $N_n(l)$ is maximum if $l = \lceil \frac{n}{2} \rceil$. Furthermore, it is known that the rounded expected linear complexity of true binary random sequences with n identically distributed symbols equals $\lceil \frac{n}{2} \rceil$. This serves as a posteriori justification for the consideration of the class of sequences that have linear complexity profiles which are bounded symmetrically in respect to $\lceil \frac{m}{2} \rceil$ for $2 \leq m \leq n$.

The linear complexity of the class of sequences as they are defined by expression (1) is lowerbounded by $\max(0, \lceil \frac{m}{2} \rceil - \lfloor K \cdot \log_2 m \rfloor)$ for length m. Analogously, the linear complexity is upperbounded by $\min(m, \lceil \frac{m}{2} \rceil + \lfloor K \cdot \log_2 m \rfloor)$. As long as $\max(0, \lceil \frac{m}{2} \rceil - \lfloor K \cdot \log_2 m \rfloor) = 0$ which is equivalent to $\lceil \frac{m}{2} \rceil \leq \lfloor K \cdot \log_2 m \rfloor$ we have obviously $\min(m, \lceil \frac{m}{2} \rceil + \lfloor K \cdot \log_2 m \rfloor) = m$.

Denote by j the maximum sequence length for which $\max(0, \lceil \frac{m}{2} \rceil - \lfloor K \cdot \log_2 m \rfloor) = 0$ is fulfilled for every m with $2 \leq m \leq j \leq n$. If $j = n$, then the number of sequences with linear complexity profile as defined in expression (1) is given by $\sum_{i=0}^{n} N_n(i) = 2^n$, which is the number of all binary sequences with length n.

III. Combinatorial Analysis and Results

In order to consider binary sequences with good linear complexity profiles for given K and for arbitrary n, it is necessary to analyze the behaviour of linear complexity profiles for $j \leq m \leq n$. By definition of j, we obtain that $\lceil \frac{j+1}{2} \rceil - \lfloor K \cdot \log_2(j+1) \rfloor) > 0$.

Since $K \cdot \log_2 x$ is monotonously increasing over the real numbers and $\lceil \frac{j+1}{2} \rceil - \lfloor K \cdot \log_2(j+1) \rfloor) > 0$ as well as $\lceil \frac{j}{2} \rceil - \lfloor K \cdot \log_2 j \rfloor = 0$, we have $\lfloor K \cdot \log_2 j \rfloor = \lfloor K \cdot \log_2(j+1) \rfloor$.

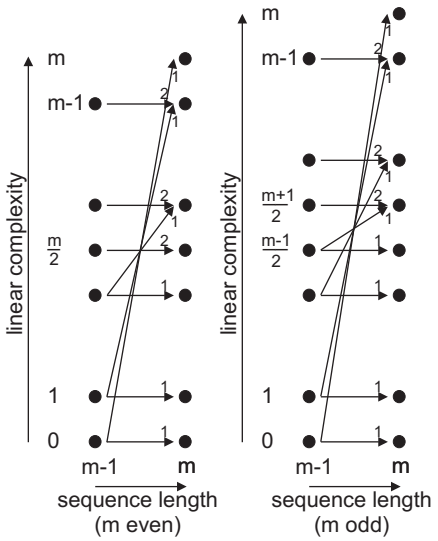

Fig. 1: Linear complexity for step from $m-1$ to m

Therefore, it yields $\lceil \frac{j+1}{2} \rceil = \lceil \frac{j}{2} \rceil + 1$ and thus, j has to be even, as it is shown in figure 2.

Since the second derivative of $K \cdot \log_2 x$ is negative and the fact that $\lfloor K \cdot \log_2 j \rfloor = \lfloor K \cdot \log_2(j+1) \rfloor$, it results that $\lfloor K \cdot \log_2(m+1) \rfloor - \lfloor K \cdot \log_2 m \rfloor \leq 1$ for $m \geq j$. Thus, $\lfloor K \cdot \log_2 m \rfloor$ can not increase more than 1 if m grows by 1 for $m \geq j$.

In the following, we consider exclusively those sequences of length n where $n > j$. Denote by R the number of jumps in $\lfloor K \cdot \log_2 m \rfloor$ for $j < m \leq n$. Then, R is obtained by

$$R = \lfloor K \cdot \log_2 n \rfloor - \lfloor K \cdot \log_2 j \rfloor. \quad (4)$$

If $R > 0$, then denote by m_i for $i = 1, \ldots, R$ the positions of the jumps in $\lfloor K \cdot \log_2 m \rfloor$, and let $j < m_1 < \ldots < m_R \leq n$.

Definition 1 *Let $m \geq j$. $Q(m)$ is defined as the $(2\lfloor K \cdot \log_2 m \rfloor + 1)$-dimensional vector whose i^{th} component represents the number of sequences of length m with linear complexity $L(s^m) = \lceil \frac{m}{2} \rceil - \lfloor K \cdot \log_2 m \rfloor - 1 + i$ for $i = 1, \ldots, 2\lfloor K \cdot \log_2 m \rfloor + 1$ and good linear complexity profile.*

If $m = j$, we have $2\lfloor K \cdot \log_2 j \rfloor = j$, and thus, we obtain the $(j+1)$-dimensional initialization vector

$$Q(j) = [N_j(0), \ldots, N_j(j)]^T. \quad (5)$$

In the following, we show how $Q(m)$ can be obtained from $Q(m-1)$ exploiting the presented properties of linear complexity shown in figure 1

- if $\lfloor K \cdot \log_2 m \rfloor > \lfloor K \cdot \log_2(m-1) \rfloor$ (subsection III.A), and
- if $\lfloor K \cdot \log_2 m \rfloor = \lfloor K \cdot \log_2(m-1) \rfloor$ (subsection III.B).

For efficiency reasons in computation, we also introduce a relation to obtain $Q(m)$ from $Q(m-2)$ for m even, which is applicable if $\lfloor K \cdot \log_2 m \rfloor = \lfloor K \cdot \log_2(m-2) \rfloor$.

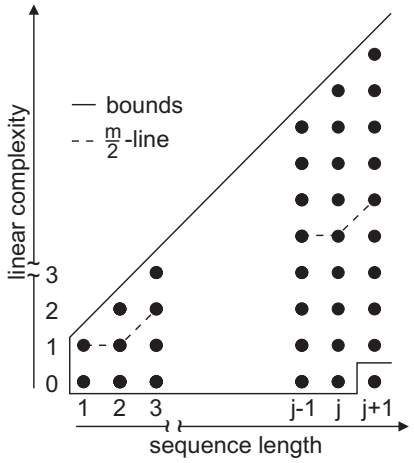

Fig. 2: Range of linear complexity profiles for $1 \leq m \leq j+1$

A. Increasing $\lfloor K \cdot \log_2 m \rfloor$

In this subsection, we will show how the number of binary sequences changes at the step from $m-1$ to m if $m \in \{m_1, \ldots, m_R\}$. Note that $2 \leq j < m_1 < \ldots < m_R$. To do this, we have to distinguish the cases

- m even, and
- m odd

for $\lfloor K \cdot \log_2 m \rfloor \geq 2$. Note that the case $\lfloor K \cdot \log_2 m \rfloor = 1$ does not exist for $K \geq 1$ and $m \in \{m_1, \ldots, m_R\}$ with $2 \leq j < m_1 < \ldots < m_R$.

I. case: m even
Here, $Q(m-1)$ is a $(2\lfloor K \cdot \log_2 m \rfloor - 1)$-dimensional vector; $Q(m)$ has dimension $2\lfloor K \cdot \log_2 m \rfloor + 1$. In this case, we have the following relations between the components of $Q(m-1)$ and $Q(m)$ that follow by the properties of linear complexity (see figure 1):

$$
\begin{aligned}
&Q_1(m) = 0, \\
&\text{for } i = 2, \ldots, \lfloor K \cdot \log_2 m \rfloor \\
&\quad Q_i(m) = Q_{i-1}(m-1), \\
&\quad Q_{\lfloor K \cdot \log_2 m \rfloor + 1}(m) = 2Q_{\lfloor K \cdot \log_2 m \rfloor}(m-1), \\
&\text{for } i = \lfloor K \cdot \log_2 m \rfloor + 2, \ldots, 2\lfloor K \cdot \log_2 m \rfloor \\
&\quad Q_i(m) = 2Q_{i-1}(m-1) + \\
&\qquad\qquad + Q_{2\lfloor K \cdot \log_2 m \rfloor + 1 - i}(m-1), \\
&\quad Q_{2\lfloor K \cdot \log_2 m \rfloor + 1}(m) = 0.
\end{aligned}
\tag{6}
$$

This relation can be modelled by a $(2\lfloor K \cdot \log_2 m \rfloor + 1) \times (2\lfloor K \cdot \log_2 m \rfloor - 1)$-matrix $M^{(e)}(\lfloor K \cdot \log_2 m \rfloor)$

$$
Q(m) \quad = \quad M^{(e)}(\lfloor K \cdot \log_2 m \rfloor) \cdot Q(m-1), \tag{7}
$$

$$
M^{(e)}(\lfloor K \cdot \log_2 m \rfloor) = \begin{bmatrix}
0 & & & & & & 0 \\
1 & & & & & & \\
0 & . & & & & & \\
 & & 1 & & & & \\
 & & 0 & 2 & & & \\
 & & 1 & 0 & 2 & & \\
0 & . & & & & . & 0 \\
1 & & & & & & 2 \\
0 & & & & & & 0
\end{bmatrix} \tag{8}
$$

□

II. case: m odd:
In this case, we also try to establish a mapping from the $(2\lfloor K \cdot \log_2 m \rfloor - 1)$-dimensional vector $Q(m-1)$ to the $(2\lfloor K \cdot \log_2 m \rfloor + 1)$-dimensional vector $Q(m)$. To do this, we also exploit the properties of linear complexity as we did before. We find

$$
\begin{aligned}
&\text{for } i = 1, \ldots, \lfloor K \cdot \log_2 m \rfloor \\
&\quad Q_i(m) = Q_i(m-1), \\
&\text{for } i = \lfloor K \cdot \log_2 m \rfloor + 1, \ldots, 2\lfloor K \cdot \log_2 m \rfloor - 1 \\
&\quad Q_i(m) = 2Q_i(m-1) + \\
&\qquad\qquad + Q_{2\lfloor K \cdot \log_2 m \rfloor + 1 - i}(m-1), \\
&\quad Q_{2\lfloor K \cdot \log_2 m \rfloor}(m) = Q_1(m-1), \\
&\quad Q_{2\lfloor K \cdot \log_2 m \rfloor + 1}(m) = 0.
\end{aligned}
\tag{9}
$$

These four subcases can also be summarized in a $(2\lfloor K \cdot \log_2 m \rfloor + 1) \times (2\lfloor K \cdot \log_2 m \rfloor - 1)$-matrix $M^{(o)}(\lfloor K \cdot \log_2 m \rfloor)$. We obtain

$$
Q(m) \quad = \quad M^{(o)}(\lfloor K \cdot \log_2 m \rfloor) \cdot Q(m-1), \tag{10}
$$

$$
M^{(o)}(\lfloor K \cdot \log_2 m \rfloor) = \begin{bmatrix}
1 & 0 & & & & & 0 \\
0 & . & & & & & \\
 & & 1 & & & & \\
 & & & 1 & & & \\
 & & & 1 & 2 & & \\
 & & . & & & . & 0 \\
0 & 1 & & & & & 2 \\
1 & & & & & & 0 \\
0 & & & & & & 0
\end{bmatrix} \tag{11}
$$

□

To summarize the results of this subsection, we have shown how $Q(m)$ can be calculated from $Q(m-1)$ for $m > j$ in case of increasing $\lfloor K \cdot \log_2 m \rfloor$ by application of expression (7) if m is even, or expression (10) if m is odd respectively.

B. Constant $\lfloor K \cdot \log_2 m \rfloor$

In this subsection, we will analyze the behaviour of $Q(m)$ at steps from $m-1$ to m, $m > j$, if $\lfloor K \cdot \log_2 (m-1) \rfloor = \lfloor K \cdot \log_2 m \rfloor$. Here, we also have to consider two cases

- $\lfloor K \cdot \log_2 m \rfloor = 1$,
- $\lfloor K \cdot \log_2 m \rfloor > 1$ with subcases m even and m odd.

I. case: $\lfloor K \cdot \log_2 m \rfloor = 1$
in this case, we can restrict our consideration to just one sequence length. Because of $K \geq 1$, the only possible sequence length $m > j$ with $\lfloor K \cdot \log_2 m \rfloor = 1$ is $m = 3$ odd. Thus, we do not have to consider the subcase m even, here. Consider the 3-dimensional vectors $Q(m-1)$ and $Q(m)$. Using the properties of linear complexity, we obtain

$$
\begin{aligned}
&Q_1(m) = Q_2(m-1), \\
&Q_2(m) = 2Q_3(m-1) + Q_2(m-1), \\
&Q_3(m) = Q_1(m-1).
\end{aligned}
\tag{12}
$$

The relation between $Q(m)$ and $Q(m-1)$ can be described by a (3×3)-matrix $G^{(o)}(\lfloor K \cdot \log_2 m \rfloor) = G^{(o)}(1)$

$$Q(m) = G^{(o)}(1) \cdot Q(m-1), \tag{13}$$

$$G^{(o)}(1) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}. \tag{14}$$

□

II. case: $\lfloor K \cdot \log_2 m \rfloor > 1$
Consider the $(2\lfloor K \cdot \log_2 m \rfloor + 1)$-dimensional vectors $Q(m)$ and $Q(m-1)$. To present the relations between the components of $Q(m)$ and $Q(m-1)$, we have to distinguish the two subcases m even and m odd.
Subcase m even:
If m is even we have the dependencies

$$\begin{aligned} &\text{for } i = 1, \ldots, \lfloor K \cdot \log_2 m \rfloor \\ &\quad Q_i(m) = Q_i(m-1), \\ &\quad Q_{\lfloor K \cdot \log_2 m \rfloor + 1}(m) = 2Q_{\lfloor K \cdot \log_2 m \rfloor + 1}(m-1), \\ &\text{for } i = \lfloor K \cdot \log_2 m \rfloor + 2, \ldots, 2\lfloor K \cdot \log_2 m \rfloor + 1 \\ &\quad Q_i(m) = 2Q_i(m-1) + \\ &\qquad\qquad + Q_{2\lfloor K \cdot \log_2 m \rfloor + 2 - i}(m-1), \end{aligned} \tag{15}$$

that can be written as a mapping using a $(2\lfloor K \cdot \log_2 m \rfloor + 1) \times (2\lfloor K \cdot \log_2 m \rfloor + 1)$-matrix $G^{(e)}(\lfloor K \cdot \log_2 m \rfloor)$

$$Q(m) \quad = \quad G^{(e)}(\lfloor K \cdot \log_2 m \rfloor) \cdot Q(m-1), \tag{16}$$

$$G^{(e)}(\lfloor K \cdot \log_2 m \rfloor) = \begin{bmatrix} 1 & 0 & & & & & 0 \\ 0 & . & & & & & \\ & & 1 & & & & \\ & & 0 & 2 & & & \\ & & 1 & 0 & 2 & & \\ 0 & . & & & & . & 0 \\ 1 & 0 & & & & 0 & 2 \end{bmatrix} \tag{17}$$

Subcase m odd:
If m is odd we obtain

$$\begin{aligned} &\text{for } i = 1, \ldots, \lfloor K \cdot \log_2 m \rfloor \\ &\quad Q_i(m) = Q_{i+1}(m-1), \\ &\text{for } i = \lfloor K \cdot \log_2 m \rfloor, \ldots, 2\lfloor K \cdot \log_2 m \rfloor \\ &\quad Q_i(m) = 2Q_{i+1}(m-1) + \\ &\qquad\qquad + Q_{2\lfloor K \cdot \log_2 m \rfloor + 2 - i}(m-1), \\ &\quad Q_{2\lfloor K \cdot \log_2 m \rfloor + 1}(m) = 2Q_1(m-1), \end{aligned} \tag{18}$$

which will be represented by means of the $(2\lfloor K \cdot \log_2 m \rfloor + 1) \times (2\lfloor K \cdot \log_2 m \rfloor + 1)$-matrix $G^{(o)}(\lfloor K \cdot \log_2 m \rfloor)$

$$Q(m) \quad = \quad G^{(o)}(\lfloor K \cdot \log_2 m \rfloor) \cdot Q(m-1), \tag{19}$$

$$G^{(o)}(\lfloor K \cdot \log_2 m \rfloor) = \begin{bmatrix} 0 & 1 & 0 & & & & 0 \\ & & . & & & & \\ & & 0 & 1 & & & \\ & & 0 & 1 & 2 & & \\ & & . & 0 & 0 & . & 0 \\ 0 & . & & & & & 2 \\ 1 & 0 & & & & & 0 \end{bmatrix}. \tag{20}$$

□

By alternating use of matrices $G^{(e)}(\lfloor K \cdot \log_2 m_i \rfloor)$ and $G^{(o)}(\lfloor K \cdot \log_2 m_i \rfloor)$ an appropriate number of times, one can determine $Q(m_{i+1}-1)$ from $Q(m_i)$. Note that $\lfloor K \cdot \log_2 m \rfloor$ is constant for $m_i \leq m \leq m_{i+1} - 1$. Since matrix multiplication is not commutative, the sequence of the matrices is not allowed to be changed in the determination of $Q(m_{i+1} - 1)$. In order to obtain more effectivity, we now analyze the behaviour of linear complexity profiles for steps from sequence length $m-2$ to m for m even and constant $\lfloor K \cdot \log_2 m \rfloor$ over this range. This means, that we need at least three subsequent sequence lengths $m-2$, $m-1$ and m for which $\lfloor K \cdot \log_2(m-2) \rfloor = \lfloor K \cdot \log_2(m-1) \rfloor = \lfloor K \cdot \log_2 m \rfloor$. Therefore, we only have to consider $\lfloor K \cdot \log_2 m \rfloor \geq 2$. Since $K \geq 1$, three subsequent sequence lengths with $\lfloor K \cdot \log_2 m \rfloor = 1$ do not exist.

Exploiting the properties of linear complexity, we obtain for steps from $m-2$ to m, $m \geq 6$ even, and $\lfloor K \cdot \log_2 m \rfloor \geq 2$

$$\begin{aligned} &\text{for } i = 1, \ldots, \lfloor K \cdot \log_2 m \rfloor \\ &\quad Q_i(m) = Q_i(m-2), \\ &\quad Q_{\lfloor K \cdot \log_2 m \rfloor + 1}(m) = 2Q_{\lfloor K \cdot \log_2 m \rfloor + 1}(m-2) + \\ &\qquad\qquad + 4Q_{\lfloor K \cdot \log_2 m \rfloor + 2}(m-2), \\ &\text{for } i = \lfloor K \cdot \log_2 m \rfloor + 2, \ldots, 2\lfloor K \cdot \log_2 m \rfloor \\ &\quad Q_i(m) = 4Q_{i+1}(m-2) + \\ &\qquad\qquad + 2Q_{2\lfloor K \cdot \log_2 m \rfloor + 2 - i}(m-2) + \\ &\qquad\qquad + Q_{2\lfloor K \cdot \log_2 m \rfloor + 3 - i}(m-2), \\ &\quad Q_{2\lfloor K \cdot \log_2 m \rfloor + 1}(m) = 2Q_1(m-2) + \\ &\qquad\qquad + Q_2(m-2), \end{aligned} \tag{21}$$

that can be described using a $(2\lfloor K \cdot \log_2 m \rfloor + 1) \times (2\lfloor K \cdot \log_2 m \rfloor + 1)$-matrix $H(\lfloor K \cdot \log_2 m \rfloor)$ by

$$Q(m) \quad = \quad H(\lfloor K \cdot \log_2 m \rfloor) \cdot Q(m-2), \tag{22}$$

$$H(\lfloor K \cdot \log_2 m \rfloor) = \begin{bmatrix} 0 & 1 & 0 & & & & 0 \\ & & . & & & & \\ & & 0 & 1 & & & \\ & & 0 & 2 & 4 & & \\ & & . & 1 & 0 & . & 0 \\ 0 & . & . & & & & 4 \\ 2 & 1 & 0 & & & & 0 \end{bmatrix}. \tag{23}$$

Now, we have the means to determine $Q(n)$ effectively applying matrices $M^{(e)}$, $M^{(o)}$, $G^{(e)}$, $G^{(o)}$ and H to $Q(j) = [N_j(0), \ldots, N_j(j)]^T$.

C. Application of the Matrices

In the previous subsection, we showed how $Q(n)$ can be obtained from $Q(j)$. The previous results can be summarized in the following theorem in order to give the number $A_K(n)$ of binary sequences of length n with good linear complexity profile.

Theorem 1 *The number $A_K(n)$ of finite binary sequences with length n and linear complexity $\lceil \frac{m}{2} \rceil - K \cdot \log_2 m \leq L(s^m) \leq \lceil \frac{m}{2} \rceil + K \cdot \log_2 m$ for $K \geq 1$ and $2 \leq m \leq n$ is given by $A_K(n) = \sum_{i=1}^{2\lfloor K \cdot \log_2 n \rfloor + 1} Q_i(n)$.*

Example Consider the set of all sequences that are given with $K = 1.2$ and $n = 12$. We obtain $j = 2$. Thus, $R = \lfloor K \cdot \log_2 n \rfloor - \lfloor K \cdot \log_2 j \rfloor = \lfloor 1.2 \cdot \log_2 12 \rfloor - \lfloor 1.2 \cdot \log_2 2 \rfloor = 4 - 1 = 3$. This means that there are three sequence lengths m for $2 < m \leq n$ where $\lfloor K \cdot \log_2 m \rfloor$ increases. These are $m_1 = 4$, $m_2 = 6$, and $m_3 = 11$. By expressions (3) and (5), we obtain $Q(2) = [1, 2, 1]^T$. Application of the matrices yields

$$\begin{aligned} Q(3) &= G^{(o)}(1) \cdot Q(2) = \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix} \cdot [1, 2, 1]^T = \\ &= [2, 4, 1]^T \end{aligned}$$

$$\begin{aligned} Q(4) &= M^{(e)}(2) \cdot Q(3) = \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \cdot [2, 4, 1]^T = \\ &= [0, 2, 8, 4, 0]^T \end{aligned}$$

$$\begin{aligned} Q(5) &= G^{(o)}(2) \cdot Q(4) = \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot [0, 2, 8, 4, 0]^T = \\ &= [2, 8, 16, 2, 0]^T \end{aligned}$$

$$\begin{aligned} Q(6) &= M^{(e)}(3) \cdot Q(5) = \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 8 \\ 16 \\ 2 \\ 0 \end{bmatrix} = \\ &= [0, 2, 8, 32, 12, 2, 0]^T \end{aligned}$$

$$\begin{aligned} Q(10) &= H^2(3) \cdot Q(6) = \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 4 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 4 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 4 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^2 \cdot \begin{bmatrix} 0 \\ 2 \\ 8 \\ 32 \\ 12 \\ 2 \\ 0 \end{bmatrix} = \\ &= [8, 32, 112, 448, 224, 56, 12]^T \end{aligned}$$

$$\begin{aligned} Q(11) &= M^{(o)}(4) \cdot Q(10) = \\ &= [8, 32, 112, 448, 896, 224, 56, 8, 0]^T \end{aligned}$$

$$\begin{aligned} Q(12) &= G^{(e)}(4) \cdot Q(11) = \\ &= [8, 32, 112, 448, 1792, 896, 224, 48, 8]^T \end{aligned}$$

By this result, we get $A_{1.2}(12) = \sum_{i=1}^{9} Q_i(12) = 3568$. □

Table I shows $A_K(n)$ for some values of n and K. One can recognize that $A_K(n)$ does nearly not differ significantly if $K \geq 1.5$.

Another interesting measure that can be easily determined by $A_K(n)$ is the conditional probability that a sequence taken out of the set of all finite binary sequences with n elements has good linear complexity profile as it was defined by (1) for given K. It is obtained as

$$P(|\lceil \frac{m}{2} \rceil - L(s^m)| \leq K \log_2 m \text{ for } m = 2 \ldots n \,|n) = \\ = \frac{A_K(n)}{2^n} \tag{24}$$

TABLE I
$A_K(n)$ for given n and K

	n					
K	20	50	100	200	500	1000
1.0	888448	9.46e14	1.06e30	1.35e60	2.74e150	8.96e300
1.1	906496	9.69e14	1.09e30	1.38e60	2.81e150	9.21e300
1.2	910464	6.76e14	1.10e30	1.39e60	2.83e150	9.27e301
1.5	1045952	1.12e15	1.26e30	1.60e60	3.27e150	1.07e301
2.0	1048540	1.13e15	1.28e30	1.61e60	3.27e150	1.07e301
3.0	1048576	1.13e15	1.28e30	1.61e60	3.27e150	1.07e301

These probabilities are shown in figures 3-6. The figures show, that for sufficient large sequence length ($n > 100$) the conditional probability in (24) only depends on K. One recognizes that for such sequence lengths, the conditional probability to find a finite binary sequence with good linear complexity profile is nearly constant for specific K. Furthermore, the conditional probability grows rapidly with increasing K.

IV. Conclusion

In this paper, we presented a new method to determine the number of finite binary sequences with good linear complexity profile. The number $A_K(n)$ of those sequences depends on the length n and the constant K which describes the range of the considered linear complexity profiles. The measure $A_K(n)$ allows the determination of the probability that an arbitrary chosen finite sequence with length n has good linear complexity profile. Finally, we show how this probabiltiy depends on n and K.

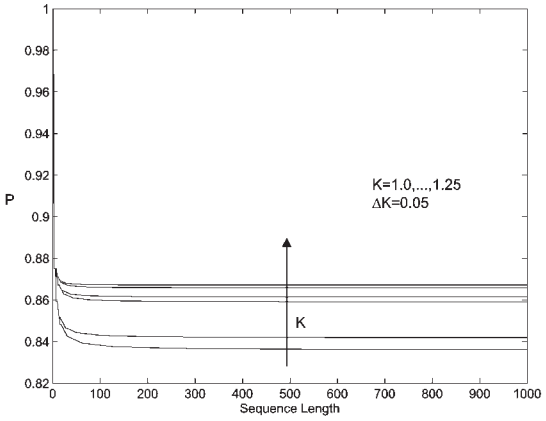

Fig. 3: Conditional Probability for n and $K = 1.0, \ldots, 1.25$

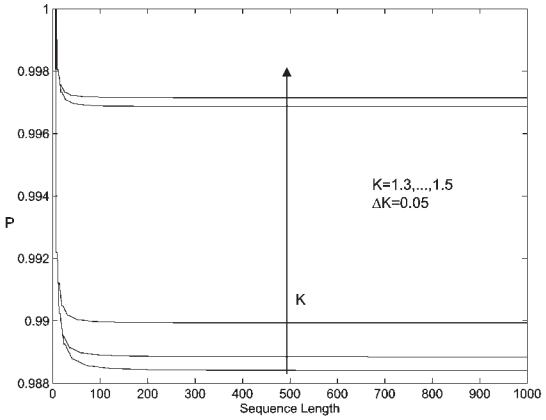

Fig. 4: Conditional Probability for n and $K = 1.3, \ldots, 1.5$

V. Acknowledgement

We are grateful to Professor Firoz Kaderali for the supervision of our work.

VI. References

[1] J.L. Massey: Shift-Register Synthesis and BCH Decoding, IEEE Trans. on Information Theory, 15 (1969)

[2] H. Niederreiter: Keystream Sequences with a Good Linear Complexity Profile for Every Starting Point, EUROCRYPT 89, Proceedings, LNCS 434, Springer Verlag, 1990

[3] R.A. Rueppel: Analysis and Design of Stream Ciphers, Springer-Verlag, Berlin 1986

[4] M. Schneider: A Note on Linear Complexity Profiles, ISCTA 97, Proceedings, 1997

[5] M. Wang: Cryptographic Aspects of Sequence Complexity Measures, Dissertation ETH, 1988

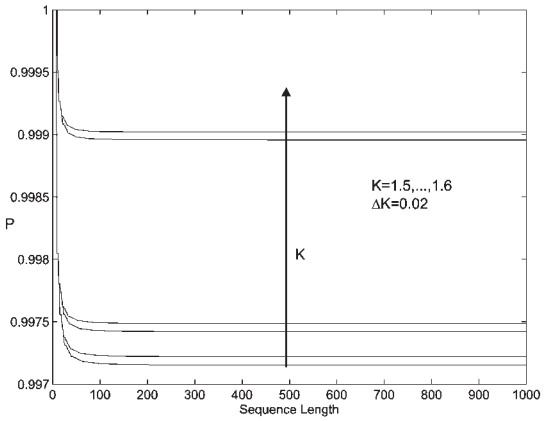

Fig. 5: Conditional Probability for n and $K = 1.5, \ldots, 1.6$

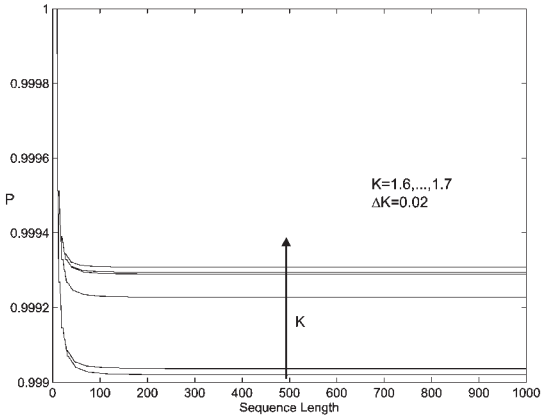

Fig. 6: Conditional Probability for n and $K = 1.6, \ldots, 1.7$