The Linear Complexity Profile and the Jump Complexity of Keystream Sequences

HARALD NIEDERREITER

Institute for Information Processing

Austrian Academy of Sciences

Sonnenfelsgasse 19, A–1010 Vienna, Austria

Abstract

We study the linear complexity profile and the jump complexity of keystream sequences in arbitrary finite fields. We solve counting problems connected with the jump complexity, establish formulas for the expected value and the variance of the jump complexity, and prove probabilistic theorems on the jump complexity profile of random sequences. We also extend earlier work on frequency distributions in the linear complexity profile to joint frequency distributions.

1. Introduction

An important tool for the assessment of keystream sequences in the context of stream ciphers is the linear complexity profile introduced by Rueppel [9], [10, Ch. 4]. Let F_q be the finite field with q elements, where q is an arbitrary prime power, let n be a positive integer, and let S be a finite or infinite sequence $s_1, s_2, \ldots$ of elements of F_q which contains at least n terms. Then the *nth linear complexity* $L_n(S)$ is defined to be the least integer k such that $s_1, s_2, \ldots, s_n$ form the first n terms of a kth-order linear feedback shift register (LFSR) sequence, where we observe the convention that the zero sequence is viewed as an LFSR sequence of order 0. The sequence $L_1(S), L_2(S), \ldots$, extended as long as $L_n(S)$ is defined, is called the *linear complexity profile* of S. We note that we have $0 \leq L_n(S) \leq n$ and $L_n(S) \leq L_{n+1}(S)$ as long as $L_n(S)$ and $L_{n+1}(S)$ are defined. Detailed studies of the linear complexity profile were carried out in [1], [5], [6], [7], [8], [9], [10, Ch. 4], [11], [12].

A new way of looking at the linear complexity profile was recently introduced by Wang [11], [12]. If S and n are as above, then the *nth jump complexity* $P_n(S)$ is defined as the number of positive integers among $L_1(S), L_2(S) - L_1(S), \ldots, L_n(S) - L_{n-1}(S)$. Thus $P_n(S)$ is the number of "jumps" in the first n terms of the linear complexity profile of S. Wang [11], [12] proved some combinatorial and elementary statistical results on the jump complexity in the special case $q = 2$, and related results were also shown by Carter [1]. In a sense, the quantity $P_n(S)$ already appeared in Niederreiter [6], since

I.B. Damgard (Ed.): Advances in Cryptology - EUROCRYPT '90, LNCS 473, pp. 174-188, 1991.

it is clear in view of [6, Lemma 1] that the quantity $j(n, S)$ defined in the proof of [6, Theorem 11] is identical with $P_n(S)$. In analogy with the terminology introduced in the previous paragraph, we call the sequence $P_1(S), P_2(S), \ldots$, extended as long as $P_n(S)$ is defined, the *jump complexity profile* of S.

The main aims of the present paper are to study the jump complexity in greater detail, to solve counting problems connected with the jump complexity, to prove probabilistic theorems on the jump complexity profile of random sequences, and to derive all these results for arbitrary q. We also extend the work in [6, Sect. 6] on frequency distributions in the linear complexity profile to joint frequency distributions.

2. Enumeration Formulas for the Jump Complexity

It is clear that we have $0 \leq P_n(S) \leq L_n(S) \leq n$ for any S containing at least n terms. We obtain a stronger upper bound from the following well-known lemma (see [4], [10, p. 34]) which holds if S contains at least $n+1$ terms.

Lemma 1. If $L_n(S) > n/2$, then $L_{n+1}(S) = L_n(S)$. If $L_n(S) \leq n/2$, then $L_{n+1}(S) = L_n(S)$ for exactly one choice of $s_{n+1} \in F_q$ and $L_{n+1}(S) = n + 1 - L_n(S)$ for exactly $q - 1$ choices of $s_{n+1} \in F_q$.

Lemma 2. $P_n(S) \leq \min(L_n(S), n - L_n(S) + 1)$.

Proof. We already noted that $P_n(S) \leq L_n(S)$, so it suffices to show $P_n(S) \leq n - L_n(S) + 1$. Consider the last jump in the first n terms of the linear complexity profile of S, say $L_m(S) < L_{m+1}(S) = \ldots = L_n(S)$. By Lemma 1 we must have $L_{m+1}(S) = m + 1 - L_m(S)$, therefore

$$P_n(S) - 1 = P_m(S) \leq L_m(S) = m + 1 - L_n(S) \leq n - L_n(S). \quad \square$$

We now establish an explicit formula for the number $N_n(L, r)$ of sequences S of elements of F_q with fixed length n and with prescribed values $L_n(S) = L$ and $P_n(S) = r$ for the nth linear complexity and the nth jump complexity, respectively. Such a formula was earlier shown by Carter [1] for $q = 2$ and Wang [11], [12] for $q = 2$ and $L = \lceil n/2 \rceil$.

Theorem 1. Let n be a positive integer and let L and r be integers with $0 \leq L, r \leq n$. Then:

(i) $N_n(L, r) = 1$ if $L = r = 0$;

(ii) $N_n(L, r) = 0$ if $L \geq 1$ and $r = 0$;

(iii) $N_n(L, r) = 0$ if $r > \min(L, n - L + 1)$;

(iv) $N_n(L, r) = \binom{\min(L-1, n-L)}{r-1} (q-1)^r q^{\min(L, n-L)}$ if $1 \leq r \leq \min(L, n - L + 1)$.

Proof. (i) is valid since $L_n(S) = P_n(S) = 0$ holds exactly for the sequence S of n zeros. (ii) holds since $L_n(S) \geq 1$ implies $P_n(S) \geq 1$. (iii) follows from Lemma 2. To prove (iv), we proceed by induction on n. The case $n = 1$ is checked immediately. Suppose the formula is shown for length n, and now consider length $n+1$. We take $1 \leq r \leq \min(L, n-L+2)$ with $1 \leq L \leq n+1$, where L and r are the prescribed values of $L_{n+1}(S)$ and $P_{n+1}(S)$, respectively. If $L \leq n/2$, then Lemma 1 implies $N_{n+1}(L,r) = N_n(L,r)$, and the induction hypothesis yields the desired formula. If $L = (n+1)/2$, so that n is odd, then we must have $L_n(S) = (n+1)/2$ by Lemma 1, and so by induction hypothesis

$$N_{n+1}(L,r) = qN_n\left(\frac{n+1}{2}, r\right) = \binom{(n-1)/2}{r-1}(q-1)^r q^{1+(n-1)/2}$$
$$= \binom{\min(L-1, n+1-L)}{r-1}(q-1)^r q^{\min(L,n+1-L)}.$$

If $L \geq (n+2)/2$, then by Lemma 1 we either have $L_n(S) = L, P_n(S) = r$, or $L_n(S) = n+1-L, P_n(S) = r-1$. Together with the induction hypothesis we get

$$N_{n+1}(L,r) = qN_n(L,r) + (q-1)N_n(n+1-L, r-1)$$
$$= \binom{n-L}{r-1}(q-1)^r q^{n-L+1} + \binom{n-L}{r-2}(q-1)^r q^{n+1-L} = \binom{n+1-L}{r-1}(q-1)^r q^{n+1-L}$$
$$= \binom{\min(L-1, n+1-L)}{r-1}(q-1)^r q^{\min(L,n+1-L)}. \square$$

From Theorem 1 we get the following alternative proof of the formula of Gustavson [2] for the number $M_n(L)$ of sequences S of elements of F_q with fixed length n and $L_n(S) = L$.

Corollary 1. $M_n(0) = 1$ and $M_n(L) = (q-1)q^{\min(2L-1, 2n-2L)}$ for $1 \leq L \leq n$.

Proof. We have $M_n(0) = N_n(0,0) = 1$. For $1 \leq L \leq n$ we get by Theorem 1,

$$M_n(L) = \sum_{r=1}^{\min(L, n-L+1)} \binom{\min(L-1, n-L)}{r-1}(q-1)^r q^{\min(L,n-L)}$$
$$= (q-1)q^{\min(L,n-L)} \sum_{r=0}^{\min(L-1,n-L)} \binom{\min(L-1, n-L)}{r}(q-1)^r$$
$$= (q-1)q^{\min(L,n-L)}(1+(q-1))^{\min(L-1,n-L)}$$
$$= (q-1)q^{\min(L,n-L)+\min(L-1,n-L)} = (q-1)q^{\min(2L-1,2n-2L)}. \square$$

Theorem 1 also yields a formula for the number $N_n(r)$ of sequences S of elements of F_q with fixed length n and $P_n(S) = r$. The special case $q = 2$ was treated by Carter [1].

Theorem 2. We have $N_n(0) = 1$ and $N_n(r) = 0$ for $r > \lceil n/2 \rceil$. For $1 \leq r \leq \lceil n/2 \rceil$ we have

$$N_n(r) = (q+1)(q-1)^r \sum_{L=r-1}^{(n-2)/2} \binom{L}{r-1} q^L \quad \text{for even} \quad n,$$

$$N_n(r) = (q+1)(q-1)^r \sum_{L=r-1}^{(n-1)/2} \binom{L}{r-1} q^L - \binom{(n-1)/2}{r-1}(q-1)^r q^{(n+1)/2} \quad \text{for odd} \quad n.$$

Proof. $N_n(0) = 1$ follows from Theorem 1 (i), (ii). If $r > \lceil n/2 \rceil$, then $r > \min(L, n-L+1)$ for $0 \leq L \leq n$, hence $N_n(r) = 0$ for $r > \lceil n/2 \rceil$ by Theorem 1 (iii). If $1 \leq r \leq \lceil n/2 \rceil$, then by Theorem 1 we have $N_n(L,r) > 0$ if and only if $r \leq \min(L, n-L+1)$, i.e. if and only if $r \leq L \leq n+1-r$. Therefore

$$N_n(r) = \sum_{L=r}^{n+1-r} N_n(L,r).$$

Now let n be even. Then by Theorem 1 (iv),

$$\begin{aligned} N_n(r) &= \sum_{L=r}^{n/2} \binom{L-1}{r-1}(q-1)^r q^L + \sum_{L=(n+2)/2}^{n+1-r} \binom{n-L}{r-1}(q-1)^r q^{n-L} \\ &= q(q-1)^r \sum_{L=r-1}^{(n-2)/2} \binom{L}{r-1} q^L + (q-1)^r \sum_{L=r-1}^{(n-2)/2} \binom{L}{r-1} q^L \\ &= (q+1)(q-1)^r \sum_{L=r-1}^{(n-2)/2} \binom{L}{r-1} q^L. \end{aligned}$$

For odd n we obtain by Theorem 1 (iv),

$$\begin{aligned} N_n(r) = &\sum_{L=r}^{(n-1)/2} \binom{L-1}{r-1}(q-1)^r q^L + \binom{(n-1)/2}{r-1}(q-1)^r q^{(n-1)/2} + \\ &+ \sum_{L=(n+3)/2}^{n+1-r} \binom{n-L}{r-1}(q-1)^r q^{n-L} \\ = q(q-1)^r &\sum_{L=r-1}^{(n-3)/2} \binom{L}{r-1} q^L + \binom{(n-1)/2}{r-1}(q-1)^r q^{(n-1)/2} + (q-1)^r \sum_{L=r-1}^{(n-3)/2} \binom{L}{r-1} q^L \\ = (q+1)(q-1)^r &\sum_{L=r-1}^{(n-1)/2} \binom{L}{r-1} q^L - \binom{(n-1)/2}{r-1}(q-1)^r q^{(n+1)/2}. \square \end{aligned}$$

3. Expected Value and Variance of the Jump Complexity

We can view $P_n = P_n(S)$ as a random variable on the space of all sequences of elements of F_q with fixed length n, where each such sequence is equiprobable (i.e., has probability q^{-n}). In the following two theorems we extend the formulas for the expected value and

the variance of P_n given by Carter [1] and Wang [11], [12] for $q = 2$ to the general case of arbitrary q.

Theorem 3. The expected value $E(P_n)$ of P_n is given by

$$E(P_n) = \frac{(q-1)n}{2q} + \frac{(q+1)^2 - (-1)^n(q-1)^2}{4(q^2+q)} - \frac{1}{(q+1)q^n}.$$

Proof. With the notation in Theorem 2 we have

$$E(P_n) = q^{-n} \sum_{r=0}^{n} r N_n(r) = q^{-n} \sum_{r=1}^{\lceil n/2 \rceil} r N_n(r).$$

If n is even, then by Theorem 2 we get

$$\begin{aligned} E(P_n) &= (q+1)q^{-n} \sum_{r=1}^{n/2} r(q-1)^r \sum_{L=r-1}^{(n-2)/2} \binom{L}{r-1} q^L \\ &= (q+1)q^{-n} \sum_{L=0}^{(n-2)/2} q^L \sum_{r=1}^{L+1} \binom{L}{r-1} r(q-1)^r \\ &= (q^2-1)q^{-n} \sum_{L=0}^{(n-2)/2} q^L \sum_{r=0}^{L} \binom{L}{r} (r+1)(q-1)^r. \end{aligned}$$

To treat the inner sum, we differentiate the identity $\sum_{r=0}^{L} \binom{L}{r} z^{r+1} = z(z+1)^L$ with respect to z and then put $z = q-1$ to obtain

$$\sum_{r=0}^{L} \binom{L}{r} (r+1)(q-1)^r = q^L + (q-1)Lq^{L-1}. \tag{1}$$

This yields

$$\begin{aligned} E(P_n) &= (q^2-1)q^{-n} \sum_{L=0}^{(n-2)/2} \left(q^{2L} + \frac{q-1}{q} L q^{2L} \right) \\ &= q^{-n}(q^n - 1) + (q^2-1)(q-1)q^{-n-1} \sum_{L=0}^{(n-2)/2} L q^{2L}. \end{aligned}$$

For any integer $k \geq 1$, differentiation of $\sum_{L=0}^{k-1} z^L = (z^k - 1)/(z-1)$ with respect to z and then multiplication by z yields

$$\sum_{L=0}^{k-1} L z^L = \frac{(k-1)z^{k+1} - kz^k + z}{(z-1)^2} \quad \text{for} \quad z \neq 1. \tag{2}$$

Putting $k = n/2, z = q^2$ in (2) we get

$$E(P_n) = 1 - q^{-n} + \frac{q}{q+1} q^{-n} \left(\frac{n-2}{2} q^n - \frac{n}{2} q^{n-2} + 1 \right)$$
$$= \frac{(q-1)n}{2q} + \frac{1}{q+1} \left(1 - q^{-n}\right)$$

by simple algebraic manipulations. For odd n we use Theorem 2 to obtain

$$E(P_n) = (q+1)q^{-n} \sum_{L=0}^{(n-1)/2} q^L \sum_{r=1}^{L+1} \binom{L}{r-1} r(q-1)^r - q^{(1-n)/2} \sum_{r=1}^{(n+1)/2} \binom{(n-1)/2}{r-1} r(q-1)^r$$

$$= (q^2-1)q^{-n} \sum_{L=0}^{(n-1)/2} q^L \sum_{r=0}^{L} \binom{L}{r} (r+1)(q-1)^r$$

$$-(q-1)q^{(1-n)/2} \sum_{r=0}^{(n-1)/2} \binom{(n-1)/2}{r} (r+1)(q-1)^r.$$

We apply (1) and get

$$E(P_n) = (q^2-1)q^{-n} \sum_{L=0}^{(n-1)/2} \left(q^{2L} + \frac{q-1}{q} L q^{2L} \right)$$

$$-(q-1)q^{(1-n)/2} \left(q^{(n-1)/2} + (q-1)\frac{n-1}{2} q^{(n-3)/2} \right)$$

$$= q^{-n}\left(q^{n+1} - 1\right) + (q^2-1)(q-1)q^{-n-1} \sum_{L=0}^{(n-1)/2} L q^{2L} - q + 1 - \frac{(q-1)^2(n-1)}{2q}.$$

Putting $k = (n+1)/2, z = q^2$ in (2) we obtain

$$E(P_n) = 1 - q^{-n} + \frac{q}{q+1} q^{-n} \left(\frac{n-1}{2} q^{n+1} - \frac{n+1}{2} q^{n-1} + 1 \right) - \frac{(q-1)^2(n-1)}{2q}$$
$$= \frac{(q-1)n}{2q} + \frac{q^2+1}{2(q^2+q)} - \frac{1}{(q+1)q^n}$$

by simple algebraic manipulations. □

Theorem 4. The variance $\mathrm{Var}(P_n)$ of P_n is given by

$$\mathrm{Var}(P_n) = \frac{(q-1)n}{2q^2} - \frac{q}{(q+1)^2} + \frac{(q-1)n}{(q+1)q^{n+1}} + \frac{1}{(q+1)q^n} - \frac{1}{(q+1)^2 q^{2n}} \quad \text{for even} \quad n,$$

$$\mathrm{Var}(P_n) = \frac{(q-1)n}{2q^2} + \frac{q^3 - 5q^2 + q + 1}{2q^2(q+1)^2} + \frac{(q-1)n}{(q+1)q^{n+1}} + \frac{2q^2 - q + 1}{(q+1)^2 q^{n+1}} - \frac{1}{(q+1)^2 q^{2n}} \quad \text{for}$$

odd n.

Proof. With the notation in Theorem 2 we have

$$E(P_n^2) = q^{-n} \sum_{r=0}^{n} r^2 N_n(r) = q^{-n} \sum_{r=1}^{\lceil n/2 \rceil} r^2 N_n(r).$$

Using Theorem 2 and the same techniques as in the proof of Theorem 3 we get

$$E(P_n^2) = \frac{(q-1)^2 n^2}{4q^2} + \frac{(3q^2-2q-1)n}{2q^2(q+1)} - \frac{q-1}{(q+1)^2} + \frac{q-1}{(q+1)^2 q^n} \quad \text{for even} \quad n,$$

$$E(P_n^2) = \frac{(q-1)^2 n^2}{4q^2} + \frac{(q^3+q-2)n}{2q^2(q+1)} + \frac{q^4+2q^3-8q^2+2q+3}{4q^2(q+1)^2} + \frac{q-1}{(q+1)^2 q^n} \quad \text{for odd} \quad n.$$

Now $\mathrm{Var}(P_n) = E(P_n^2) - E(P_n)^2$, and so an application of Theorem 3 yields the desired formulas by straightforward algebraic manipulations. □

4. The Jump Complexity Profile of Random Sequences

In this section we study the behavior of the jump complexity profile for random infinite sequences. Note that by Theorem 3 we expect $P_n(S)$ to be close to $(q-1)n/(2q)$ and that by Theorem 4 there should be deviations that are at least of the order of magnitude $n^{1/2}$. We set up a suitable probabilistic model as in earlier probabilistic studies (see [6], [7], [8]). Let F_q^∞ be the set of all infinite sequences of elements of F_q. A probability measure h on F_q^∞ is defined by first considering the uniform probability measure μ on F_q which assigns the measure $1/q$ to each element of F_q and then letting h be the complete product measure on F_q^∞ induced by μ. We say that a property π of sequences $S \in F_q^\infty$ holds *with probability 1* if the set of all $S \in F_q^\infty$ which have the property π has h-measure 1.

Theorem 5. If f is a nonnegative function on the positive integers such that $\sum_{n=1}^{\infty} n e^{-qf(n)^2/n} < \infty$, then with probability 1 we have

$$|P_n(S) - \frac{(q-1)n}{2q}| \le f(n) \quad \text{for all sufficiently large} \quad n.$$

Proof. Choose $c \in (0, \frac{1}{2})$ such that

$$\frac{2c}{3(1-2c)^3} - c + \frac{c}{(q-1)^2} = \frac{1}{2(q-1)}. \tag{3}$$

This is possible since the left-hand side of (3) tends to 0 as $c \to 0$ and it tends to ∞ as $c \to \frac{1}{2}$ from the left. Put

$$F(n) = \min\left(f(n), \frac{cn}{q} - 1\right), \tag{4}$$

where we assume that n is so large that $F(n) \geq \frac{1}{2}$ (note that the condition on f implies $\lim_{n\to\infty} f(n) = \infty$). For fixed n consider

$$D_n = \left\{S \in F_q^\infty : P_n(S) > \frac{(q-1)n}{2q} + F(n)\right\}.$$

Let $k(n)$ be the least integer satisfying

$$k(n) > \frac{(q-1)n}{2q} + F(n). \tag{5}$$

For even n we get by Theorem 2,

$$h(D_n) = q^{-n} \sum_{r=k(n)}^{n/2} N_n(r) = (q+1)q^{-n} \sum_{r=k(n)}^{n/2} (q-1)^r \sum_{L=r-1}^{(n-2)/2} \binom{L}{r-1} q^L.$$

For the inner sum we have

$$\sum_{L=r-1}^{(n-2)/2} \binom{L}{r-1} q^L \leq q^{(n-2)/2} \sum_{L=r-1}^{(n-2)/2} \binom{L}{r-1} = q^{(n-2)/2} \binom{n/2}{r},$$

thus

$$h(D_n) < 2q^{-n/2} \sum_{r=k(n)}^{n/2} \binom{n/2}{r} (q-1)^r.$$

Since $r \geq k(n) > (q-1)n/(2q)$, the terms of the last sum form a decreasing function of r, hence

$$h(D_n) < 2q^{-n/2} \left(\frac{n}{2} - k(n) + 1\right) \binom{n/2}{k(n)} (q-1)^{k(n)} \leq nw \tag{6}$$

with

$$w = \binom{n/2}{k(n)} (q-1)^{k(n)} q^{-n/2}.$$

We use Stirling's formula in the form

$$\log(n!) = \left(n + \frac{1}{2}\right) \log n - n + O(1).$$

Then

$$\begin{aligned}\log w =& \frac{n+1}{2}\log\frac{n}{2} - \left(k(n)+\frac{1}{2}\right)\log k(n) - \left(\frac{n+1}{2}-k(n)\right)\log\left(\frac{n}{2}-k(n)\right)\\ &+ k(n)\log(q-1) - \frac{n}{2}\log q + O(1)\\ \le& \frac{n}{2}\log\frac{n}{2} - k(n)\log k(n) - \left(\frac{n}{2}-k(n)\right)\log\left(\frac{n}{2}-k(n)\right) + k(n)\log(q-1) - \frac{n}{2}\log q + O(1)\\ &=: u + O(1).\end{aligned}$$

Put

$$d(n) = k(n) - \frac{(q-1)n}{2q}.$$

Then by straightforward manipulations

$$u = -\left(\frac{n}{2q}-d(n)\right)\log\left(1-\frac{2qd(n)}{n}\right) - \left(\frac{(q-1)n}{2q}+d(n)\right)\log\left(1+\frac{2qd(n)}{(q-1)n}\right).$$

Note that by the definitions of $F(n)$ and $k(n)$ in (4) and (5) we have $0 \le d(n) \le cn/q$, hence $0 \le 2qd(n)/n \le 2c$. For $0 \le z \le 2c$ we have

$$\log(1-z) \ge -z - \frac{z^2}{2} - \frac{z^3}{3(1-2c)^3}$$

by Taylor's theorem. Using also $\log(1+z) \ge z - \frac{1}{2}z^2$ for $z \ge 0$, we get

$$\begin{aligned}u \le& \left(\frac{n}{2q}-d(n)\right)\left(\frac{2qd(n)}{n} + \frac{2q^2d(n)^2}{n^2} + \frac{8q^3d(n)^3}{3(1-2c)^3n^3}\right)\\ &- \left(\frac{(q-1)n}{2q}+d(n)\right)\left(\frac{2qd(n)}{(q-1)n} - \frac{2q^2d(n)^2}{(q-1)^2n^2}\right)\\ <& -\frac{q^2d(n)^2}{(q-1)n} + \frac{2q^2d(n)^3}{n^2}\left(\frac{2}{3(1-2c)^3} - 1 + \frac{1}{(q-1)^2}\right)\\ =& -\frac{q^2d(n)^2}{(q-1)n} + \frac{q^2d(n)^3}{c(q-1)n^2} = \frac{q^2d(n)^2}{(q-1)n}\left(-1+\frac{d(n)}{cn}\right) \le -\frac{qd(n)^2}{n}\end{aligned}$$

by (3) and $d(n) \le cn/q$. It follows that

$$\log w < -\frac{qd(n)^2}{n} + O(1),$$

and so (6) and $d(n) > F(n)$ yield

$$h(D_n) = O\left(ne^{-qF(n)^2/n}\right). \tag{7}$$

For odd n we get by Theorem 2,

$$h(D_n) = q^{-n}\sum_{r=k(n)}^{(n+1)/2} N_n(r) \le (q+1)q^{-n}\sum_{r=k(n)}^{(n+1)/2}(q-1)^r\sum_{L=r-1}^{(n-1)/2}\binom{L}{r-1}q^L.$$

Proceeding as in the case of even n, we obtain

$$h(D_n) < \left(q^{1/2}+1\right) q^{-n/2} \sum_{r=k(n)}^{(n+1)/2} \binom{(n+1)/2}{r} (q-1)^r.$$

Since $r \geq k(n) > (q-1)n/(2q) + 1/2$, the terms of the last sum form a decreasing function of r, hence

$$h(D_n) < \left(q^{1/2}+1\right) q^{-n/2} \left(\frac{n+1}{2} - k(n) + 1\right) \binom{(n+1)/2}{k(n)} (q-1)^{k(n)} < C(q) n w_1 \quad (8)$$

with $C(q)$ depending only on q and

$$w_1 = \binom{(n+1)/2}{k(n)} (q-1)^{k(n)} q^{-n/2}.$$

By Stirling's formula,

$$\begin{aligned} \log w_1 =& \frac{n+2}{2} \log \frac{n+1}{2} - \left(k(n) + \frac{1}{2}\right) \log k(n) - \left(\frac{n+2}{2} - k(n)\right) \log \left(\frac{n+1}{2} - k(n)\right) \\ &+ k(n) \log(q-1) - \frac{n}{2} \log q + O(1) \\ \leq& u + O(1), \end{aligned}$$

where u is as before. Using (8) and the bound $u < -qd(n)^2/n$ shown above, we see that (7) also holds for odd n. From (4) we deduce

$$n\, e^{-qF(n)^2/n} < n\, e^{-qf(n)^2/n} + n\, e^{1-(c^2 n/q)},$$

and so the hypothesis on f implies $\sum_{n=1}^{\infty} n\, e^{-qF(n)^2/n} < \infty$. In view of (7) this shows that $\sum_{n=1}^{\infty} h(D_n) < \infty$. An application of the Borel–Cantelli lemma [3, p. 228] yields that the set of all S for which $S \in D_n$ for infinitely many n has h-measure 0. In other words, with probability 1 we have $S \in D_n$ for at most finitely many n. From the definition of D_n it follows then that with probability 1 we have

$$P_n(S) \leq \frac{(q-1)n}{2q} + F(n) \leq \frac{(q-1)n}{2q} + f(n) \quad \text{for all sufficiently large} \quad n, \quad (9)$$

where we applied (4) in the second inequality.

To obtain a corresponding lower bound, we proceed by similar arguments. Choose $c_1 \in \left(0, \frac{1}{2}\right)$ such that

$$\frac{2c_1}{3(1-2c_1)^3(q-1)} - \frac{c_1}{q-1} + (q-1)c_1 = \frac{1}{2(q-1)}. \quad (10)$$

Put

$$G(n) = \min\left(f(n), \frac{(q-1)c_1 n}{q} - 1\right), \quad (11)$$

where we assume that n is so large that $G(n) \geq 0$. For fixed n consider

$$E_n = \left\{ S \in F_q^\infty : P_n(S) < \frac{(q-1)n}{2q} - G(n) \right\}.$$

Let $m(n)$ be the largest integer satisfying

$$m(n) < \frac{(q-1)n}{2q} - G(n). \tag{12}$$

For even n we get by Theorem 2,

$$h(E_n) = q^{-n} \sum_{r=0}^{m(n)} N_n(r) = q^{-n} + (q+1)q^{-n} \sum_{r=1}^{m(n)} (q-1)^r \sum_{L=r-1}^{(n-2)/2} \binom{L}{r-1} q^L.$$

By treating the inner sum as before we obtain

$$h(E_n) < q^{-n} + 2q^{-n/2} \sum_{r=1}^{m(n)} \binom{n/2}{r} (q-1)^r.$$

Since $r \leq m(n) < (q-1)n/(2q)$, the terms of the last sum form an increasing function of r, hence

$$h(E_n) < 3q^{-n/2} m(n) \binom{n/2}{m(n)} (q-1)^{m(n)} < \frac{3n}{2} w_2 \tag{13}$$

with

$$w_2 = \binom{n/2}{m(n)} (q-1)^{m(n)} q^{-n/2}.$$

Put

$$e(n) = \frac{(q-1)n}{2q} - m(n).$$

Then as before

$$\log w_2 \leq u_1 + O(1)$$

with

$$u_1 = -\left(\frac{n}{2q} + e(n)\right) \log\left(1 + \frac{2qe(n)}{n}\right) - \left(\frac{(q-1)n}{2q} - e(n)\right) \log\left(1 - \frac{2qe(n)}{(q-1)n}\right).$$

By the definitions of $G(n)$ and $m(n)$ in (11) and (12) we have $0 \leq e(n) \leq (q-1)c_1 n/q$, hence $0 \leq 2qe(n)/(q-1)n \leq 2c_1$. Using the same lower bound for $\log(1-z)$ as before and also $\log(1+z) \geq z - \frac{1}{2}z^2$ for $z \geq 0$, we get

$$\begin{aligned}
u_1 \leq & -\left(\frac{n}{2q} + e(n)\right)\left(\frac{2qe(n)}{n} - \frac{2q^2e(n)^2}{n^2}\right) + \\
& + \left(\frac{(q-1)n}{2q} - e(n)\right)\left(\frac{2qe(n)}{(q-1)n} + \frac{2q^2e(n)^2}{(q-1)^2n^2} + \frac{8q^3e(n)^3}{3(1-2c_1)^3(q-1)^3n^3}\right) \\
< & -\frac{q^2e(n)^2}{(q-1)n} + \frac{2q^2e(n)^3}{n^2}\left(\frac{2}{3(1-2c_1)^3(q-1)^2} - \frac{1}{(q-1)^2} + 1\right) \\
= & -\frac{q^2e(n)^2}{(q-1)n} + \frac{q^2e(n)^3}{(q-1)^2c_1n^2} = \frac{q^2e(n)^2}{(q-1)n}\left(-1 + \frac{e(n)}{(q-1)c_1n}\right) \leq -\frac{qe(n)^2}{n}
\end{aligned}$$

by (10) and $e(n) \leq (q-1)c_1 n/q$. Therefore

$$\log w_2 < -\frac{qe(n)^2}{n} + O(1),$$

and so (13) and $e(n) > G(n)$ yield

$$h(E_n) = O\left(n\, e^{-qG(n)^2/n}\right).$$

It can again be proved that this bound also holds for odd n. Together with the hypothesis on f this shows that $\sum_{n=1}^{\infty} h(E_n) < \infty$. By applying the Borel–Cantelli lemma as before we deduce that with probability 1 we have

$$P_n(S) \geq \frac{(q-1)n}{2q} - G(n) \geq \frac{(q-1)n}{2q} - f(n) \quad \text{for all sufficiently large} \quad n.$$

Together with (9) this yields the result of the theorem. □

Theorem 6. With probability 1 we have

$$\overline{\lim_{n\to\infty}} (n \log n)^{-1/2} |P_n(S) - \frac{(q-1)n}{2q}| \leq \left(\frac{2}{q}\right)^{1/2}.$$

Proof. For a positive integer m consider the function

$$f(n) = \left(\frac{2+m^{-1}}{q}\right)^{1/2} (n \log n)^{1/2}.$$

Then

$$n\, e^{-q\, f(n)^2/n} = n^{-1-m^{-1}},$$

and so $\sum_{n=1}^{\infty} n\, e^{-qf(n)^2/n} < \infty$. Thus Theorem 5 shows that with probability 1 we have

$$|P_n(S) - \frac{(q-1)n}{2q}| \leq \left(\frac{2+m^{-1}}{q}\right)^{1/2} (n \log n)^{1/2} \quad \text{for all sufficiently large} \quad n.$$

This property holds simultaneously for all m with probability 1 since the countable intersection of sets of h-measure 1 has again h-measure 1. The desired conclusion follows. □

Corollary 2. With probability 1 we have

$$\lim_{n\to\infty} \frac{P_n(S)}{n} = \frac{q-1}{2q}.$$

We note that the result of Corollary 2 was already shown in [6, eq.(6)] by using the deeper methods of that paper. It suffices to observe, as we already did in Section 1 of the present paper, that the quantity $j(n, S)$ in [6] is identical with $P_n(S)$.

5. Joint Frequency Distributions in the Linear Complexity Profile

In [6, Sect. 6] we studied frequency distributions in the linear complexity profile by considering the quantity $Z(N; c; S)$ defined as follows. For any integers c and N with $N \geq 1$ and any $S \in F_q^\infty$, $Z(N; c; S)$ is given as the number of integers n with $1 \leq n \leq N$ for which $L_n(S) = (n + c)/2$. It was shown in [6, Theorem 11] that with probability 1 we have

$$\lim_{N\to\infty} \frac{Z(N; c; S)}{N} = \frac{1}{2}(q-1)q^{-|c-(1/2)|-(1/2)} \quad \text{for all integers} \quad c. \tag{14}$$

We now extend this work to joint frequency distributions. For $S \in F_q^\infty$ and integers c_0, c_1, N with $N \geq 1$ let $Z(N; c_0, c_1; S)$ be the number of integers n with $1 \leq n \leq N$ for which $L_n(S) = (n + c_0)/2$ and $L_{n+1}(S) = (n + c_1)/2$. Then we are interested in the existence of

$$\lim_{N\to\infty} \frac{Z(N; c_0, c_1; S)}{N}.$$

Since $L_n(S) \leq L_{n+1}(S)$, we can assume that $c_0 \leq c_1$, for otherwise we have the trivial case where $Z(N; c_0, c_1; S) = 0$ for all N and S. We have to distinguish the cases $c_0 = c_1$ and $c_0 < c_1$.

Theorem 7. With probability 1 we have

$$\lim_{N\to\infty} \frac{Z(N; c, c; S)}{N} = \frac{1}{2}(q-1)q^{-|c-1|-1} \quad \text{for all integers} \quad c.$$

Proof. First let $c \geq 1$. If $L_n(S) = (n + c)/2$, then $L_n(S) > n/2$, and so by Lemma 1 we have $L_{n+1}(S) = (n + c)/2$. Therefore $Z(N; c, c; S) = Z(N; c; S)$, and the desired result follows from (14). Now let $c \leq 0$. Suppose $n, 1 \leq n \leq N$, is such that $L_n(S) = L_{n+1}(S) = (n + c)/2$. We go to the next jump, say

$$\frac{n+c}{2} = L_n(S) = L_{n+1}(S) = \ldots = L_{n+k}(S) < L_{n+k+1}(S),$$

where $k \geq 1$. Then by Lemma 1,

$$L_{n+k+1}(S) = n + k + 1 - L_{n+k}(S) = \frac{n + 2k + 2 - c}{2}.$$

It follows that for $1 \leq i \leq k$ we have

$$L_{n+k+i}(S) \geq L_{n+k+1}(S) = \frac{n + 2k + 2 - c}{2} \geq \frac{n + 2k + 2}{2} > \frac{n + k + i}{2},$$

and so Lemma 1 yields $L_{n+2k}(S) = L_{n+k+1}(S) = (n+2k+2-c)/2$. If $n+2k \le N$, this shows that $n+2k$ counts towards $Z(N; 2-c; S)$. If $n+2k > N$, then n is the largest value in $[1, N]$ which counts towards $Z(N; c, c; S)$. Therefore

$$Z(N; c, c; S) \le Z(N; 2-c; S) + 1. \tag{15}$$

Now suppose $n, 1 \le n \le N$, is such that $L_n(S) = (n+2-c)/2$, where again $c \le 0$. We go to the previous jump, say

$$L_{n-k}(S) < L_{n-k+1}(S) = \ldots = L_n(S) = \frac{n+2-c}{2},$$

where $k \ge 1$. Then by Lemma 1,

$$L_{n-k}(S) = n - k + 1 - L_{n-k+1}(S) = \frac{n-2k+c}{2}.$$

We claim that

$$L_{n-2k+i}(S) = L_{n-k}(S) \quad \text{for} \quad 0 \le i \le k. \tag{16}$$

For suppose that for some $i, 0 \le i < k$, we had $L_{n-2k+i}(S) < L_{n-2k+i+1}(S) = L_{n-k}(S)$. Then Lemma 1 yields

$$\begin{aligned} L_{n-2k+i}(S) &= n - 2k + i + 1 - L_{n-2k+i+1}(S) = n - 2k + i + 1 - \frac{n-2k+c}{2} \\ &= \frac{n-2k-c}{2} + i + 1 > \frac{n-2k+i}{2}. \end{aligned}$$

Another application of Lemma 1 implies $L_{n-2k+i+1}(S) = L_{n-2k+i}(S)$, a contradiction. Thus (16) is shown. Using (16) with $i = 0, 1$, we get $L_{n-2k}(S) = L_{n-2k+1}(S) = (n-2k+c)/2$. If $n-2k \ge 1$, this shows that $n-2k$ counts towards $Z(N; c, c; S)$. If $n-2k < 1$, then n is the smallest value in $[1, N]$ which counts towards $Z(N; 2-c; S)$. Thus

$$Z(N; 2-c; S) \le Z(N; c, c; S) + 1. \tag{17}$$

It follows from (14), (15), and (17) that if $c \le 0$, then with probability 1 we have

$$\lim_{N\to\infty} \frac{Z(N; c, c; S)}{N} = \lim_{N\to\infty} \frac{Z(N; 2-c; S)}{N} = \frac{1}{2}(q-1)q^{c-2} = \frac{1}{2}(q-1)q^{-|c-1|-1}. \square$$

Now let $c_0 < c_1$. If $L_n(S) = (n+c_0)/2$ and $L_{n+1}(S) = (n+c_1)/2$, then it follows from Lemma 1 that $L_{n+1}(S) = n + 1 - L_n(S)$, and so we must have $c_1 = 2 - c_0$.

Theorem 8. If $c_0 < c_1 = 2 - c_0$, then with probability 1 we have

$$\lim_{N\to\infty} \frac{Z(N; c_0, c_1; S)}{N} = \frac{1}{2}(q-1)^2 q^{-c_1}.$$

If $c_0 < c_1 \ne 2 - c_0$, then $Z(N; c_0, c_1; S) = 0$ for all N and S.

Proof. The second part was shown above. Now let $c_0 < c_1 = 2 - c_0$ and note that this implies $c_1 \geq 2$. From Lemma 1 we see that $L_n(S) = (n + c_0)/2$ and $L_{n+1}(S) = (n + c_1)/2$ hold simultaneously if and only if $L_{n+1}(S) - L_n(S) = c_1 - 1$. Therefore $Z(N; c_0, c_1; S) = J(N; c_1 - 1; S)$, where the latter denotes the number of integers n with $1 \leq n \leq N$ and $L_{n+1}(S) - L_n(S) = c_1 - 1$. Thus with probability 1 we have

$$\lim_{N\to\infty} \frac{Z(N; c_0, c_1; S)}{N} = \lim_{N\to\infty} \frac{J(N; c_1 - 1; S)}{P_N(S)} \cdot \frac{P_N(S)}{N}$$

$$= (q-1)q^{1-c_1}\frac{q-1}{2q} = \frac{1}{2}(q-1)^2 q^{-c_1},$$

where we applied a result in [6, p.208] and Corollary 2. □

Since there are just countably many pairs (c_0, c_1) of integers, it follows that with probability 1 we have the asymptotic frequency distributions in Theorems 7 and 8 simultaneously for all choices of c_0 and c_1.

References

[1] G.D. Carter: *Aspects of Local Linear Complexity*, Ph.D. thesis, Univ. of London, 1989.

[2] F.G. Gustavson: Analysis of the Berlekamp-Massey linear feedback shift-register synthesis algorithm, *IBM J. Res. Develop.* **20**, 204-212 (1976).

[3] M. Loève: *Probability Theory*, 3rd ed., Van Nostrand, New York, 1963.

[4] J.L. Massey: Shift-register synthesis and BCH decoding, *IEEE Trans. Information Theory* **15**, 122-127 (1969).

[5] H. Niederreiter: Sequences with almost perfect linear complexity profile, *Advances in Cryptology - EUROCRYPT '87*, Lecture Notes in Computer Science, Vol. **304**, pp. 37-51, Springer, Berlin, 1988.

[6] H. Niederreiter: The probabilistic theory of linear complexity, *Advances in Cryptology -EUROCRYPT '88*, Lecture Notes in Computer Science, Vol. **330**, pp. 191-209, Springer, Berlin, 1988.

[7] H. Niederreiter: A combinatorial approach to probabilistic results on the linear-complexity profile of random sequences, *J. of Cryptology*, to appear.

[8] H. Niederreiter: Keystream sequences with a good linear complexity profile for every starting point, *Advances in Cryptology - EUROCRYPT '89*, Lecture Notes in Computer Science, Springer, Berlin, to appear.

[9] R.A. Rueppel: Linear complexity and random sequences, *Advances in Cryptology - EUROCRYPT '85*, Lecture Notes in Computer Science, Vol. **219**, pp. 167-188, Springer, Berlin, 1986.

[10] R.A. Rueppel: *Analysis and Design of Stream Ciphers*, Springer, Berlin, 1986.

[11] M.Z. Wang: *Cryptographic Aspects of Sequence Complexity Measures*, Ph.D. dissertation, ETH Zürich, 1988.

[12] M.Z. Wang: Linear complexity profiles and jump complexity, submitted to *J. of Cryptology*.