An algorithm for the word entropy

Sébastien Ferenczi, Christian Mauduit, Carlos Gustavo Moreira

An algorithm for the word entropy

Sébastien Ferenczi
IMPA - CNRS UMI 2924
Estrada Dona Castorina 110, 22460-320 Rio de Janeiro, RJ, Brasil
and
Institut de Mathématiques de Marseille, UMR 7373 CNRS,
163, avenue de Luminy, 13288 Marseille Cedex 9, France

Christian Mauduit
Université d'Aix-Marseille and Institut Universitaire de France,
Institut de Mathématiques de Marseille, UMR 7373 CNRS,
163, avenue de Luminy, 13288 Marseille Cedex 9, France

Carlos Gustavo Moreira
Instituto de Matemática Pura e Aplicada,
Estrada Dona Castorina 110,
22460-320 Rio de Janeiro, RJ, Brasil

Abstract: For any infinite word w on a finite alphabet A, the complexity function p_w of w is the sequence counting, for each non-negative n, the number $p_w(n)$ of words of length n on the alphabet A that are factors of the infinite word w and the entropy of w is the quantity $E(w) = \lim_{n\to\infty} \frac{1}{n} \log p_w(n)$. For any given function f with exponential growth, Mauduit and Moreira introduced in [MM17] the notion of word entropy $E_W(f) = \sup\{E(w), w \in \mathbb{A}^{\mathbb{N}}, p_w \leq f\}$ and showed its links with fractal dimensions of sets of infinite sequences with complexity function bounded by f. The goal of this work is to give an algorithm to estimate with arbitrary precision $E_W(f)$ from finitely many values of f.

2010 Mathematics Subject Classification: 68R15, 37B10, 37B4, 28D20.

Keywords: combinatorics on words, symbolic dynamics, entropy.

This work was supported by CNPq, FAPERJ and the Agence Nationale de la Recherche project ANR-14-CE34-0009 MUDERA.

The authors wish to thank the referee for his very careful scrutiny.

1 Introduction

This work concerns the little-explored field of word combinatorics in positive entropy, which means the study of infinite words on a finite alphabet with a complexity function (see Definition 2.2) of exponential growth. There are not many results on this topic, besides the well-known one of Grillenberger [Gri73] who built symbolic systems of any given entropy.

Mauduit and Moreira introduced in [MM17] new notions in this context with the arithmetic motivation to study sets of numbers from the interval $[0, 1]$ whose expansion (in a given base q) has a complexity function bounded by a given function f. The determination of the Hausdorff dimension of these sets gave rise to a new quantity $E_W(f)$, called *word entropy* of f, which turns to be equal to the topological entropy of the shift on the set of corresponding expansions.

The computation of $E_W(f)$ is trivial when $E_0(f)$, the *exponential growth rate* of f (defined in (1)), is equal to zero or if f is itself a complexity function. Otherwise, results can be surprising, even when f is very regular: for example, in [MM17] it is shown that for the function f defined for any non-negative integer n by $f(n) = 3^{\lceil \frac{n}{2} \rceil}$, we have $E_W(f) = \log(\frac{1+\sqrt{5}}{2})$. Another striking result (see Theorem 2.9 from [MM18]) says that if f verifies the quite natural conditions $(\mathcal{C}^*)$ (see Definition 3.2), then the ratio $E_W(f)/E_0(f)$ lies always in the interval $]\frac{1}{2}, 1]$ and moreover we have

$$\inf\{\frac{E_W(f)}{E_0(f)}), f \ satisfies \ (\mathcal{C}^*)\} = \frac{1}{2}.$$

Indeed, in the overwhelming majority of cases, we do not have access to an exact value of the word entropy. Thus in this work we propose an algorithm to get an approximate value of the word entropy, using in depth the combinatorial properties of the symbolic system.

2 Definitions and notations

We denote by q a fixed integer greater or equal to 2, by A the finite alphabet $A = \{0, 1, \ldots, q-1\}$, by $A^* = \bigcup_{k\geq 0} A^k$ the set of finite words on the alphabet A and by $A^{\mathbb{N}}$ the set of infinite words (or infinite sequences of letters) on the alphabet A. More generally, if $\Sigma \subset A^*$, we denote by $\Sigma^{\mathbb{N}}$ the set of infinite words obtained by concatenating elements of Σ. If $w \in A^{\mathbb{N}}$ we denote by $L(w)$ the set of finite factors of w:

$$L(w) = \{v \in A^*,\ \exists\, (v', v'') \in A^* \times A^{\mathbb{N}},\ w = v'vv''\}$$

and, for any non-negative integer n, we write $L_n(w) = L(w) \cap A^n$. For any $Y \subset A^{\mathbb{N}}$ and $n \in \mathbb{N}$ we denote $L_n(Y) = \bigcup_{w\in Y} L_n(w)$. If $w \in A^n, n \in \mathbb{N}$ we denote $|w| = n$ the length of the word w and if S is a finite set, we denote by $|S|$ the number of elements of S. For any $(a, b) \in \mathbb{R}^2$ with $a \leq b$, we denote by $[\![a, b]\!]$ the set $[a, b] \cap \mathbb{Z}$ and for any x real number, we denote $\lfloor x \rfloor = \max\{n \in \mathbb{Z}, n \leq x\}$, $\lceil x \rceil = \min\{n \in \mathbb{Z}, x \leq n\}$ and $\{x\} = x - \lfloor x \rfloor$.

Let us recall the following classical lemma concerning sub-additive sequences due to Fekete [Fek23]:

Lemma 2.1. *If $(a_n)_{n\geq 1}$ is a sequence of real numbers such that $a_{n+n'} \leq a_n + a_{n'}$ for any positive integers n and n', then the sequence $\left(\frac{a_n}{n}\right)_{n\geq 1}$ converges to $\inf_{n\geq 1} \frac{a_n}{n}$.*

Definition 2.2. The *complexity function* of $w \in A^{\mathbb{N}}$ is defined for any non-negative integer n by $p_w(n) = |L_n(w)|$.

For any $w \in A^{\mathbb{N}}$ and for any $(n, n') \in \mathbb{N}^2$ we have $L_{n+n'}(w) \subset L_n(w)L_{n'}(w)$ so that $p_w(n+n') \leq p_w(n)p_w(n')$ and it follows from Lemma 2.1 that for any $w \in A^{\mathbb{N}}$, the sequence $\left(\frac{1}{n}\log p_w(n)\right)_{n\geq 1}$ converges to $\inf_{n\geq 1} \frac{1}{n}\log p_w(n)$. We denote

$$E(w) = \lim_{n\to\infty} \frac{1}{n}\log p_w(n) = h_{top}(X(w), T)$$

the topological entropy of the symbolic dynamical system $(X(w), T)$ where T is the one-sided shift on $A^{\mathbb{N}}$ and $X = \overline{orb_T(w)}$ is the closure of the orbit of w under the action of T in $A^{\mathbb{N}}$ ($A^{\mathbb{N}}$ is equipped with the product topology of the discrete topology on A, i.e. the topology induced by the distance $d(w, w') = \exp(-\min\{n \in \mathbb{N} |\, w_n \neq w'_n\})$).

The complexity function gives information about the statistical properties of an infinite sequence of letters. In this sense, it constitutes one possible way to measure the random behaviour of an infinite sequence: see [Que87, Fer99, PF02].

3 Exponential rate of growth and word entropy of a function

For any given function f from $\mathbb{N}$ to $\mathbb{R}^+$, we denote

$$W(f) = \{w \in A^{\mathbb{N}}, p_w(n) \leq f(n), \forall n \in \mathbb{N}\},$$

$$\mathcal{L}_n(f) = \bigcup_{w \in W(f)} L_n(w)$$

and $E_0(f)$ the limiting lower exponential growth rate of f

$$E_0(f) = \lim_{n \to \infty} \inf \frac{1}{n} \log f(n). \tag{1}$$

For any $(n, n') \in \mathbb{N}^2$ we have $\mathcal{L}_{n+n'}(f) \subset \mathcal{L}_n(f)\mathcal{L}_{n'}(f)$ so that the sequence $\left(\frac{1}{n} \log |\mathcal{L}_n(f)|\right)_{n \geq 1}$ converges to $\inf_{n \geq 1} \frac{1}{n} \log |\mathcal{L}_n(f)|$, which is the topological entropy of the subshift $(W(f), T)$:

$$h_{top}(W(f), T) = \lim_{n \to +\infty} \frac{1}{n} \log |\mathcal{L}_n(f)| = \inf_{n \geq 1} \frac{1}{n} \log |\mathcal{L}_n(f)|.$$

The notion of w-entropy (or word-entropy) of f is defined in [MM17] as follow :

Definition 3.1. If f is a function from $\mathbb{N}$ to $\mathbb{R}^+$, the *w-entropy* (or *word entropy*) of f is the quantity

$$E_W(f) = \sup_{w \in W(f)} E(w).$$

The papers [MM10] and [MM12] concern the case $E_0(f) = 0$ and [MM17] the case of positive entropy. In particular the word entropy of f is equal to the topological entropy of the subshift $(W(f), T)$ (see Theorem 2.3 from [MM17]): for any function f from $\mathbb{N}$ to $\mathbb{R}^+$, we have

$$E_W(f) = \lim_{n \to +\infty} \frac{1}{n} \log(|\mathcal{L}_n(f)|) = h_{top}(W(f), T).$$

(see also beginning of Section 4 from [MM17] and Chapter 8 from [Wal82] to understand this result as a consequence of the variational principle).

The word entropy of f allows to compute exactly the fractal dimensions of the set of real numbers from the interval $[0, 1]$ the $q-$adic expansion of which has a complexity function bounded by f. (see Theorem 5.1 from [MM17]). Note that several authors have applied the notion of dimension introduced by Hausdorff in [Hau19] to number theoretical

problems (see [Bug04, Chapters V and VI] for a very good survey on these questions and [Fal90, Chapters 2 and 3] for basic definitions concerning fractal dimensions).

Definition 3.2. We say that a function f from $\mathbb{N}$ to $\mathbb{R}^+$ satisfies the conditions $(\mathcal{C}^*)$ if
i) for any $n \in \mathbb{N}$ we have $f(n+1) > f(n) \geq n+1$;
ii) for any $(n, n') \in \mathbb{N}^2$ we have $f(n+n') \leq f(n)f(n')$.

For any function f from $\mathbb{N}$ to $\mathbb{R}^+$ we have $E_W(f) \leq E_0(f)$ and it is easy to give examples of function f for which the entropy ratio $E_W(f)/E_0(f)$ can be made arbitrarily small (see beginning of Section 7 from [MM17]). When f satisfies the quite natural conditions $(\mathcal{C}^*)$, it still might happen that $E_W(f) < E_0(f)$ (see Sections 7.2 and 7.4 from [MM17]), but Mauduit and Moreira proved the following theorem (see Theorem 4.2 and Remark 4.3 from [MM18]):

Theorem 3.3. *If f is a function from $\mathbb{N}$ to $\mathbb{R}^+$ satisfying the conditions $(\mathcal{C}^*)$, then $E_W(f) > \frac{1}{2}E_0(f)$.*

Moreover, the constant $\frac{1}{2}$ in Theorem 3.3 is optimal (see Theorem 5.1 from [MM18]).

As mentioned in [MM17] it is in general more difficult to compute $E_W(f)$ than $E_0(f)$. The goal of this work is to give an algorithm which allows us to estimate with arbitrary precision $E_W(f)$ from finitely many values of f, if we know already $E_0(f)$ and have some information on the speed with which this limit is approximated.

4 The algorithm

We assume that the function f from $\mathbb{N}$ to $\mathbb{R}^+$ satisfies the conditions $(\mathcal{C}^*)$. We don't loose generality with this assumption, since if there exists an integer n such that $f(n) < n+1$ we have $E_W(f) = 0$ and if not, it follows from Remark 7.3 from [MM17] that we may always change the function f by a function $\tilde{\tilde{f}}$ satisfying conditions $(\mathcal{C}^*)$ and such that $E_W(\tilde{\tilde{f}}) = E_W(f)$.

Theorem 4.1. *There is an algorithm which gives, starting from the function f and $\varepsilon \in]0,1[$, a quantity h such that $(1-\varepsilon)h \leq E_W(f) \leq h$. The quantity h depends explicitely*

on ε, $E_0(f)$, N, $f(1)$, ..., $f(N)$, for an integer N which depends explicitely on ε, $E_0(f)$ and an integer n_0. larger than an explicit function of ε and $E_0(f)$ and such that

$$\frac{\log f(n)}{n} < (1 + \frac{E_0(f)\varepsilon}{210(4 + 2E_0(f))})E_0(f) \quad \textit{for any} \quad n \in [\![n_0, 2n_0 - 1]\!].$$

We shall now give the algorithm and prove Theorem 4.1. The function f is given and henceforth we omit to mention it in $E_0(f)$ and $E_W(f)$.

Description of the algorithm

For $\varepsilon \in]0, 1[$ given, let

$$\delta := \frac{E_0\varepsilon}{105(4 + 2E_0)} < \frac{\varepsilon}{210} \tag{2}$$

and

$$K := \lceil \delta^{-1} \rceil + 1. \tag{3}$$

We choose a positive integer

$$n_0 \geq \max(K, \frac{4K^2}{420^3 E_0}, \frac{16}{KE_0^4}) \tag{4}$$

such that for any integer $n \geq n_0$

$$\frac{\log f(n)}{n} < (1 + \frac{\delta}{2})E_0. \tag{5}$$

In view of conditions $(\mathcal{C}^*)$, this last condition is equivalent to $\frac{\log f(n)}{n} < (1 + \frac{\delta}{2})E_0$ for any $n \in [\![n_0, 2n_0 - 1]\!]$. We choose intervals which will be so large that all the lengths of words we manipulate stay in one of them. Namely, for each non-negative integer t, let

$$n_{t+1} := \exp(K((1 + \delta)^2 E_0 n_t + E_0)).$$

We take

$$N := n_K.$$

We choose now a set $Y \subset A^N$ and we define

$$q_n(Y) := |L_n(Y)|$$

for $n \in [\![1, N]\!]$. We look at those Y for which

$$q_n(Y) \leq f(n) \tag{6}$$

for any $n \in [\![1, N]\!]$ and choose one among them such that

$$\min_{1 \leq n \leq N} \frac{\log q_n(Y)}{n}$$

is maximum. Henceforth we omit to mention Y in the notation $q_n(Y)$.

Proposition 4.2. *We have*

$$\min_{1 \leq n \leq N} \frac{\log q_n}{n} \geq E_W.$$

Proof. It follows from Section 4.3 of [MM17] (see (4)) that there is $\hat{w} \in W(f)$ with $p_n(\hat{w}) \geq \exp(E_W n)$ for any positive integer n. For such a word $\hat{w}$, let

$$X := L_N(\hat{w}) \subset A^N.$$

We have, for each for $n \in [\![1, N]\!]$, $L_n(X) = L_n(\hat{w})$ and $f(n) \geq |L_n(\hat{w})| = p_n(\hat{w}) \geq \exp(E_W n)$. Thus X is one of the possible $Y$and the result follows from the maximality of $\min_{1 \leq n \leq N} \frac{\log q_n}{n}$. □

The next lemma shows that on one of the large intervals we have defined, the quantity $\frac{\log q_n}{n}$ will be almost constant:

Lemma 4.3. *There exists a non-negative integer $r < K$, such that*

$$\frac{\log q_{n_r}}{n_r} < (1+\delta)\frac{\log q_{n_{r+1}}}{n_{r+1}}.$$

Proof. Otherwise we would have

$$\frac{\log q_{n_0}}{n_0} \geq (1+\delta)^K \frac{\log q_{n_K}}{n_K}. \tag{7}$$

As $K > \frac{1}{\delta}$, we have $(1+\delta)^K = e^{K \log(1+\delta)} > e^{\frac{1}{\delta}\log(1+\delta)} > \frac{9}{4}$ for $\delta < \frac{1}{2}$. By Proposition 4.2, we have $\frac{\log q_{n_K}}{n_K} \geq E_W$, so that (7) would implies that $\frac{\log q_{n_0}}{n_0} \geq \frac{9}{4} E_W$. But it follows from (6) that $q_{n_0} \leq f(n_0)$ and from (5) that $\frac{\log q_{n_0}}{n_0} < (1+\frac{\delta}{2})E_0 \leq \frac{9}{8}E_0$ for $\delta < \frac{1}{4}$. Finally we would have $E_W \leq \frac{1}{2}E_0$ which would contradict Theorem 3.3. □

If we put

$$h := \frac{\log q_{n_r}}{n_r},$$

the next proposition follows immediately from Proposition 4.2 :

Proposition 4.4. *We have*

$$h \geq E_W.$$

We shall use the estimates given by the following lemma:

Lemma 4.5. *We have*

$$\frac{E_0}{2} \leq h \leq E_0(1+\frac{\delta}{2})$$

Proof. It follows from Proposition 4.4 and Theorem 3.3 that $h \geq E_W > \frac{E_0}{2}$. On the other hand, it follows from (6) that $q_{n_r} \leq f(n_r)$ and as $n_r > n_0$, it follows from (5) that $h \leq E_0(1+\frac{\delta}{2})$. □

What remains to prove is the following proposition (which, understandably, does not use the maximality of $\min_{1\leq n\leq N} \frac{\log q_n}{n}$).

Proposition 4.6. *We have*

$$(1-\varepsilon)h \leq E_W.$$

Proof. Our strategy is to build a word w such that, for any positive integer n,

$$\exp((1-\varepsilon)hn) \leq p_n(w) \leq f(n),$$

which gives the conclusion by definition of E_W. To build the word w, we shall define an integer m and build successive subsets of $L_m(Y)$. We order any such a subset Z (lexicographically for example) and define $w(Z)$ by using a Champernowne-type construction: namely, if $Z = \{\beta_1, \beta_2, ..., \beta_t\}$, we build the infinite word

$$w(Z) := \beta_1\beta_2 \dots \beta_t\beta_1\beta_1\beta_1\beta_2\beta_1\beta_3 \dots \beta_t\beta_t\beta_1\beta_1\beta_1 \dots \beta_t\beta_t\beta_t \dots$$

made by concatenation of all words in Z followed by the concatenations of all pairs of words of Z followed by the concatenations of all triples of words of Z, etc... (see [Ch33] and [MS98] for statistical properties of Champernowne words).

The word $w(Z)$ will satisfy $\exp((1-\varepsilon)hn) \leq p_n(w(Z))$ for any positive integer n as soon as

$$|Z| \geq \exp((1-\varepsilon)hm)$$

since, for every positive integer k, we will have at least $|Z|^k$ factors of length km in $w(Z)$.

The successive (decreasing) subsets Z of $L_m(Y)$ we build will all have cardinality at least $\exp((1-\varepsilon)hm)$ and the words $w(Z)$ will satisfy $p_n(w(Z)) \leq f(n)$ for n in an interval which will increase at each new set Z we build and ultimately contains all the integers.

We begin by an estimate on q_n using the value of h.

Lemma 4.7. *For any $n \in [\![1, N]\!]$, we have $q_n \leq \exp(hn + hn_r)$.*

Proof. For any integer non-negative integer $n \leq N$ we write $n = an_r + b$ with a non-negative integer and $b \in [\![0, n_r - 1]\!]$. As we have $L_n(Y) \subset L_{an_r}(Y)L_b(Y) \subset (L_{n_r}(Y))^a L_b(Y)$ and $q_{n_r} = \exp(hn_r)$, we get

$$q_n \leq q_{n_r}^a q_b = \exp(ahn_r)q_b \leq \exp(hn)q_{n_r} = \exp(hn_r)\exp(hn).$$

□

The following lemma uses only properties of f, independently of the definition of Y.

Lemma 4.8. *For any integer $n \geq n_0$, there exists $n' \in [\![n, (1+\delta)n]\!]$ such that*

$$f(n' + j) \geq \exp(\frac{E_0 j}{2})f(n')$$

for every positive integer j.

Proof. Otherwise there would exist j_0 such that $f(n + j_0) < \exp(\frac{E_0 j_0}{2})f(n)$, then there would exist j_1 such that $f(n + j_0 + j_1) < \exp(\frac{E_0 j_1}{2})f(n + j_0) < \exp(E_0 \frac{j_0 + j_1}{2})f(n)$ and so on until we are out of the interval. Thus we would get some integer $s > \delta n$ such that $f(n + s) < \exp(\frac{E_0 s}{2})f(n)$, but then, by the choice of n_0, we would have $f(n + s) < \exp(E_0 \frac{s}{2})\exp((1 + \frac{\delta}{2})E_0 n)$. This last quantity is smaller than $\exp(E_0(n + s))$, because $s > \delta n$ implies that $n\frac{\delta}{2} + \frac{s}{2} < s$, This would contradict the definition of E_0. □

We are now ready to begin our construction. Our first aim is to define two lengths of words, $\hat{n}$ and m, which will be in the interval $[n_r, n_{r+1}]$ but with m much larger than $\hat{n}$ and a set Z_1 of words of length m of the form $\gamma\theta$, for words γ of length $\hat{n}$, such that the word $\gamma\theta\gamma$ is in $L_{m+\hat{n}}(Y)$. Thus, for a while, we shall be interested in twin occurrences of words.

Let $\hat{n}$ be the n' of Lemma 4.8 defined for $n = Kn_r$ (note that the fact that $\hat{n}$ satisfies the conclusion of Lemma 4.8 will not be used before Lemma 4.13 much later). Let

$$\hat{N} := \lceil \exp(\frac{E_0}{2}) f(\hat{n}) \rceil$$

and

$$Y_1 := L_{\hat{N}}(Y).$$

We know that $Kn_r \leq \hat{n} \leq (1+\delta)Kn_r$. The first inequality implies that $n_r < \delta\hat{n}$ and the second inequality implies (by the initial choice of the n_r) that $\hat{N} < n_{r+1}$. We write $n_{r+1} = a\hat{N} + b$ with a positive integer and $b \in [\![0, \hat{N}-1]\!]$ and we use the defining property of r in Lemma 4.3, which translates into

$$q_{n_{r+1}} \geq \exp(\frac{hn_{r+1}}{1+\delta}) = \exp(\frac{h(a\hat{N}+b)}{1+\delta}).$$

On the other hand, we have by Lemma 4.7,

$$q_{n_{r+1}} \leq q_{\hat{N}}^a q_b \leq q_{\hat{N}}^a \exp(hb + hn_r).$$

Hence we get

$$q_{\hat{N}}^a \geq \exp(\frac{ha\hat{N} + hb}{1+\delta} - h(b + n_r))$$

and, as $a \geq 1$, this implies

$$q_{\hat{N}} \geq \exp(h(\frac{\hat{N}}{1+\delta} - \frac{\delta b}{1+\delta} - n_r)).$$

As $b < \hat{N}$ and $n_r < \delta\hat{N}$, we get

$$|Y_1| = q_{\hat{N}} > \exp((1-3\delta)h\hat{N}).$$

For the moment, we fix a word W in Y_1. The word W has $\hat{N} - \hat{n} + 1$ factors of length $\hat{n}$ and we claim that $\hat{N} - \hat{n} + 1 > (1 + \frac{E_0}{2}) f(\hat{n}) > f(\hat{n})$: because $\hat{N} > (1 + \frac{E_0}{2} + \frac{E_0^2}{8}) f(\hat{n})$, it is enough to prove $\frac{E_0^2}{8} f(\hat{n}) \geq \hat{n}$, which comes from $f(\hat{n}) \geq \exp(E_0\hat{n})$ (as item (ii) of Condition $(\mathcal{C}^*)$ ensures the limit E_0 is an infimum), $\exp(E_0\hat{n}) > \frac{(E_0\hat{n})^2}{2}$, and $\hat{n} \geq n_0 K \geq \frac{16}{E_0^4}$. There are at most $f(\hat{n})$ distinct factors of length $\hat{n}$. We make the list of the $c \leq f(\hat{n})$ different words occurring in W, the j-th one appearing a_j times, with $\sum_{j=1}^{c} a_j > \hat{N} - \hat{n}$. We look

at pairs of occurrences of the same factor, beginning at two different positions $s < t$. We denote such a pair by (s,t) and say two such pairs (s,t) and (s',t') are distinct if $t \neq t'$. Thus there are at least $\sum_{j=1}^{c}(a_j - 1) \geq \hat{N} - \hat{n} + 1 - f(\hat{n})$ distinct pairs. To each pair (s,t) we associate the interval $[s, t+\hat{n}[$. The union of these intervals contains at least $\hat{N} - \hat{n} + 1 - f(\hat{n}) + \hat{n} - 1 = \hat{N} - f(\hat{n})$ integer points.

Now we use the following elementary

Lemma 4.9. *Given a finite family of intervals $(I_j)_{1\leq j\leq d}$, there is a subfamily of disjoint intervals $(I_j)_{j\in J}$ such that*

$$\sum_{j\in J} |I_j| \geq |\cup_{i=1}^{d} I_j|.$$

Proof. We number the I_j by ascending order of their lowest elements. Let $\hat{d}$ be the largest j such that $I_j \supset I_d$. We can remove all the I_j for $\hat{d} < j \leq d$, if they exist. Then if $I_j \cap I_{j+2} \neq \emptyset$, I_{j+1} must be included in $I_j \cup I_{j+2} \cup \ldots \cup I_{\hat{d}}$ and we can remove I_{j+1}. Thus, after removing some intervals and renumbering, we can suppose all the $I_j \cap I_{j+2}$ are empty. Then either the family of even-numbered intervals or the family of odd-numbered intervals satisfies our requirements. □

We apply Lemma 4.9 to the above intervals $[s, t+\hat{n}[$, for the word W. Thus we get some ℓ and $s_1 < t_1 < \ldots < s_\ell < t_\ell$, such that the same factor of W occurs at positions s_i and t_i and the sum of the lengths $\sum_{i=1}^{\ell}(t_i + \hat{n} - s_i)$ is at least

$$\frac{\hat{N} - f(\hat{n})}{2} \geq \frac{E_0}{4 + 2E_0}\hat{N},$$

because $\hat{N} \geq e^{\frac{E_0}{2}} f(\hat{n})$ and $\frac{1-e^{-\frac{E_0}{2}}}{2} \geq \frac{\frac{E_0}{2}}{2(1+\frac{E_0}{2})}$.

Since $t_i + \hat{n} - s_i \geq \hat{n}$ for each $i \leq \ell$, we have $\ell \leq \frac{\hat{N}}{\hat{n}}$.

Now, if we look at all W in Y_1, the number of possible choices for the pairs $(s_i, t_i), 1 \leq i \leq \ell$ is at most

$$\sum_{\ell=1}^{\frac{\hat{N}}{\hat{n}}} \binom{\hat{N}}{2\ell} \leq \exp(\frac{4\hat{N}\log\hat{n}}{\hat{n}})$$

and this is smaller than $\exp(\delta h \hat{N})$ because $\hat{n} \geq n_0 K \geq K^2 > \frac{1}{\delta^2}$, thus

$$\frac{4\log\hat{n}}{\hat{n}} \leq 8\delta^2 \log\frac{1}{\delta} < \delta\frac{E_0}{2} \leq \delta h.$$

The cardinality of Y_1 is at least $\exp((1-3\delta)h\hat{N})$, thus we can find a subset $Y_2 \subset Y_1$ of at least $\exp((1-4\delta)h\hat{N})$ elements of Y_1 which have the same choice of pairs (s_i, t_i).

We define, for $(s,t) \in [\![1, \hat{N}]\!]^2$ with $s < t$, the projections $\pi_{s,t} \colon Y_2 \to A^{t-s}$ by $\pi_{s,t}(\beta_1, \beta_2, \ldots, \beta_{\hat{N}}) = (\beta_s, \beta_{s+1}, \ldots, \beta_{t-1})$.

Let

$$\tilde{\varepsilon} = \frac{\varepsilon}{15} = \frac{7(4+2E_0)\delta}{E_0} > 14\delta. \tag{8}$$

Lemma 4.10. *There is a pair (s_i, t_i) such that*

$$|\pi_{s_i, t_i+\hat{n}}(Y_2)| \geq \exp((1-\tilde{\varepsilon})h \cdot (t_i + \hat{n} - s_i)).$$

Proof. Suppose by contradiction that for each $i \in [\![1, \ell]\!]$ we have

$$|\pi_{s_i, t_i+\hat{n}}(Y_2)| < \exp((1-\tilde{\varepsilon})h \cdot (t_i + \hat{n} - s_i)).$$

The interval $[1, \hat{N}[$ can be written as the union of the intervals $[s_i, t_i + \hat{n}[, i \in [\![1, \ell]\!]$ with at most $\ell + 1$ holes. Let M be the sum of the lengths of the holes, we have proved M is at most $(1 - \frac{E_0}{4+2E_0})\hat{N}$.

By Lemma 4.7, an upper bound for the number of possible sequences in these holes is

$$\exp((\ell+1)hn_r + hM) \leq \exp(hM + 2\delta h\hat{N}).$$

This would give an upper estimate for the total number of words in Y_2 of the order of

$$\exp(h\hat{N}) \exp(-\tilde{\varepsilon}h \frac{E_0}{4+2E_0}\hat{N}) \exp(2\delta h\hat{N}) \leq$$

$$\exp(h\hat{N}) \exp(-7\delta h\hat{N}) \exp(2\delta h\hat{N}) = \exp((1-5\delta)h\hat{N}),$$

which would contradict the lower estimate $\exp((1-4\delta)h\hat{N})$. □

Now we fix a pair (s_i, t_i) such that

$$|\pi_{s_i, t_i+\hat{n}}(Y_2)| \geq \exp((1-\tilde{\varepsilon})h(t_i + \hat{n} - s_i)).$$

For a word in Y_2, the sequence of its letters whose positions are in the interval $[s_i, t_i+\hat{n}[$ is such that its last $\hat{n}$ letters coincide with its first $\hat{n}$ letters. Bounding $q_{t_i+\hat{n}-s_i}$ by Lemma 4.7 and using $n_r < \delta\hat{n}$, we get

$$|\pi_{s_i, t_i+\hat{n}}(Y_2)| \leq \exp(h(t_i + \hat{n} - s_i + \delta\hat{n} - \hat{n})),$$

which because of the above choice of the pair implies $(1-\delta)\hat{n} \leq \tilde{\varepsilon}(t_i + \hat{n} - s_i)$ and $t_i + \hat{n} - s_i - \hat{n} > \frac{1}{2\tilde{\varepsilon}}\hat{n}$.

If we put

$$m := t_i - s_i,$$

we have

$$m > \frac{\hat{n}}{2\tilde{\varepsilon}}. \tag{9}$$

We shall need the following upper bound.

Lemma 4.11. *We have* $m < \exp(\frac{E_0}{2}\tilde{\varepsilon}m)$.

Proof. Let ϕ the function defined for any $x \in \mathbb{R}^+$ by $\phi(x) = \exp(\frac{E_0}{2}\tilde{\varepsilon}x) - x$.

The function ϕ is increasing on the interval $[\frac{K^2}{2\tilde{\varepsilon}}, +\infty[$: we have $\phi'(x) = \frac{E_0}{2}\tilde{\varepsilon}\exp(\frac{E_0}{2}\tilde{\varepsilon}x) - 1$ and $\phi'(\frac{K^2}{2\tilde{\varepsilon}}) > 0$ because it follows from (2), (3) and (8) that $K^2 > (\frac{420}{E_0\varepsilon})^2 > \frac{420^2}{{E_0}^2\varepsilon} > \frac{8}{{E_0}^2\tilde{\varepsilon}}$, so that [1]

$$\exp(\frac{E_0}{4}K^2) > \frac{E_0}{4}K^2 > \frac{2}{E_0\tilde{\varepsilon}}.$$

It follows from (9) and (4) that $m > \frac{\hat{n}}{2\tilde{\varepsilon}} > \frac{Kn_0}{2\tilde{\varepsilon}} > \frac{K^2}{2\tilde{\varepsilon}}$, so that

$$\phi(m) > \phi(\frac{K^2}{2\tilde{\varepsilon}})$$

and it follows from (2), (3) and (8) that ${E_0}^3K^4 > \frac{(420)^3}{\varepsilon^3}\frac{14}{\tilde{\varepsilon}} > \frac{(420)^3 14}{\tilde{\varepsilon}} > \frac{(4)^3 3}{\tilde{\varepsilon}}$, so that [2]

$$\phi(\frac{K^2}{2\tilde{\varepsilon}}) = \exp(\frac{E_0}{4}K^2) - \frac{K^2}{2\tilde{\varepsilon}} > \frac{{E_0}^3K^6}{6(4)^3} - \frac{K^2}{2\tilde{\varepsilon}} > 0.$$

□

The set

$$Z_1 := \pi_{s_i,t_i}(Y_2)$$

is made with words of length m of the type $\gamma\theta$ for words γ of length $\hat{n}$, such that the word $\gamma\theta\gamma$ is in $\pi_{s_i,t_i+\hat{n}}(Y_2)$. Thus

$$|Z_1| = |\pi_{s_i,t_i+\hat{n}}(Y_2)| \geq \exp((1-\tilde{\varepsilon})h(m+\hat{n}).$$

[1] For any $x \in \mathbb{R}^+$, we have $\exp(x) > x$.

[2] For any $x \in \mathbb{R}^+$, we have $\exp(x) > \frac{x^3}{6}$.

Then we consider the prefixes of length $6\tilde{\varepsilon}m \geq 3\hat{n}$ of words of Z_1 and their suffixes of length $6\tilde{\varepsilon}m \geq 3\hat{n}$. By Lemma 4.7, and $n_r < \delta\hat{n}$, there are at most $\exp(12\tilde{\varepsilon}hm + 2\delta h\hat{n})$ such subwords and, by choosing those which are more frequent, we define a new set $Z_2 \subset Z_1$ in which all the words have the same prefix γ_1 of length $6\tilde{\varepsilon}m$ and all the words have the same suffix γ_2 of length $6\tilde{\varepsilon}m$, with $|Z_2| \geq |Z_1| \exp(-12\tilde{\varepsilon}hm - 2\delta h\hat{n})$ and $2\delta\hat{n} \leq (1-\tilde{\varepsilon})m$, thus

$$|Z_2| \geq \exp((1 - 13\tilde{\varepsilon})hm).$$

As a consequence of the definition of Z_2, all words of Z_2 have the same prefix of length $\hat{n}$, which is a prefix γ_0 of γ_1. As Z_2 is included in Z_1, any word of Z_2 is of the form $\gamma_0\theta$ and the word $\gamma_0\theta\gamma_0$ is in $L_{m+\hat{n}}(Y)$.

We can now reap a (small) first benefit of all this construction: by using the above property of γ_0, we can bound by below $f(n)$ the number of very short factors of $w(Z_2)$.

Claim 4.12. *We have $p_{w(Z_2)}(n) \leq f(n)$ for any $n \in [\![1, \hat{n}+1]\!]$.*

Proof. For $1 \leq n \leq \hat{n}+1$, a factor x of length n of $w(Z_2)$ either is a factor of a word of Z_2 and this word is some $\gamma_0\theta$, or else is made with a suffix of length $u \in [\![1, n-1]\!]$ of a word $\gamma_0\theta$ of Z_2 concatenated with a prefix of length $n-u \in [\![1, n-1]\!]$ of another word $\gamma_0\theta'$ of Z_2, thus x is a factor of $\gamma_0\theta\gamma_0$. In both cases x is a factor of a word in $L_{m+\hat{n}}(Y)$, thus is in $L_n(Y)$. Thus our claim is satisfied as $|L_n(Y)| \leq q_n \leq f(n)$. □

Let us shrink again our set of words.

Lemma 4.13. *For a given subset Z of Z_2, there exists $Z' \subset Z$,*

$$|Z'| \geq (1 - \exp(-(j-1)\frac{E_0}{2}))^j |Z|,$$

such that the total number of factors of length $\hat{n}+j$ of all words $\gamma_0\theta\gamma_0$ such that $\gamma_0\theta$ is in Z' is at most $f(\hat{n}+j) - j$.

Proof. Let $w_1, \ldots, w_c$, with $c \leq f(\hat{n}+j)$, the factors of length $\hat{n}+j$ of all words $\gamma_0\theta\gamma_0$ such that $\gamma_0\theta$ is in Z. If $c < f(\hat{n}+j)$, we add arbitrary words (we call them ghost factors) w_d, $c < d \leq f(\hat{n}+j)$ of length $\hat{n}+j$, to make $f(\hat{n}+j)$ different words. For such a word $\gamma_0\theta\gamma_0$, its number of factors of length $\hat{n}+j$ is at most $m + \hat{n} - (\hat{n}+j) + 1$.

The proportion of subsets $\{w_{i_1}, \dots w_{i_j}\}$ of j words (among the possible $f(\hat{n}+j)$ factors of length $\hat{n}+j$, including ghost factors) such that no w_{i_r} is a factor of $\gamma_0\theta\gamma_0$ is at least

$$\frac{\binom{f(\hat{n}+j)-(m-j+1)}{j}}{\binom{f(\hat{n}+j)}{j}} = \frac{f(\hat{n}+j)-(m-j+1)}{f(\hat{n}+j)} \cdots \frac{f(\hat{n}+j)-m}{f(\hat{n}+j)-j+1}$$

$$> (\frac{f(\hat{n}+j)-m}{f(\hat{n}+j)})^j > (1-\exp((j-1)\frac{E_0}{2}))^j$$

as $f(\hat{n}+j) \geq \exp(jE_0/2)f(\hat{n})$, by choice of $\hat{n}$ after Lemma 4.8 and $m \leq \hat{N}-\hat{n} \leq \exp(\frac{E_0}{2})f(\hat{n})$.

Thus on average a subset of j factors w_t intersects a proportion at most $1-(1-e^{(j-1)\frac{E_0}{2}})^j$ of the words $\gamma_0\theta\gamma_0$ for $\gamma_0\theta$ in Z. There are as many words $\gamma_0\theta$ in Z as corresponding words $\gamma_0\theta\gamma_0$. Thus there exists a set of j factors w_t and a subset Z' of Z of cardinality at least $(1-e^{-(j-1)\frac{E_0}{2}})^j|Z|$ such that none of the j factors w_t is a factor of a word $\gamma_0\theta\gamma_0$ for $\gamma_0\theta$ in Z'. □

We start from Z_2 and apply successively Lemma 4.13 from $j=2$ to $j=6\tilde{\varepsilon}m$, getting $6\tilde{\varepsilon}m-1$ successive sets Z'. At the end, we get a set Z_3 such that the total number of factors of length $\hat{n}+j$ of words $\gamma_0\theta\gamma_0$ for $\gamma_0\theta$ in Z_3 is at most $f(\hat{n}+j)-j$ for $j=2,\dots,6\tilde{\varepsilon}m$ and $\frac{|Z_3|}{|Z_2|}$ is at least

$$\prod_{2\leq j\leq 6\tilde{\varepsilon}m}(1-\exp(-(j-1)\frac{E_0}{2}))^j \geq \prod_{j\geq 2}(1-\exp(-(j-1)\frac{E_0}{2}))^j := p_0.$$

We have

$$\log p_0 = \sum_{j\geq 2} j\log(1-\exp(-(j-1)\frac{E_0}{2}))$$

$$> \sum_{j\geq 2} \frac{-j\exp(-(j-1)\frac{E_0}{2})}{1-\exp(-(j-1)\frac{E_0}{2})}.$$ [3]

It follows that

$$\log p_0 > \frac{-1}{1-\exp(-\frac{E_0}{2})}\sum_{j\geq 2} j\exp(-(j-1)\frac{E_0}{2}) = \frac{-\exp(-\frac{E_0}{2})(2-\exp(-\frac{E_0}{2}))}{(1-\exp(-\frac{E_0}{2}))^3} \geq \frac{-1}{(1-\exp(-\frac{E_0}{2}))^3},$$

[3] For any $x \in]0,1[$, we have $\log(1-x) > -\frac{x}{1-x}$.

which implies [4]

$$p_0 \geq \exp(-(1+\frac{2}{E_0})^3).$$

Now $(1+\frac{2}{E_0})^3$ is smaller than $\tilde{\varepsilon}hm$ because

$$h \geq \frac{E_0}{2}, \quad \tilde{\varepsilon}m \geq \frac{\hat{n}}{2} \geq \frac{Kn_0}{2}$$

and from (4)

$$K\frac{E_0}{4}n_0 \geq \frac{K^3}{420^3} \geq (1+\frac{2}{E_0})^3,$$

thus

$$|Z_3| \geq \exp((1-14\tilde{\varepsilon})hm).$$

We can now bound the number of short factors by using the factors we have just deleted and properties of the words γ_0, γ_1 and γ_2.

Claim 4.14. *We have $p_{w(Z_3)}(n) \leq f(n)$ for any $n \in [\![1, 6\tilde{\varepsilon}m]\!]$.*

Proof. Claim 4.12 is still valid for $Z_3 \subset Z_2$, so we look at a factor x in $w(Z_3)$ of length $\hat{n}+j$ with $j \in [\![2, 6\tilde{\varepsilon}m - \hat{n}]\!]$. If x is a factor of some $\gamma_0\theta\gamma_0$ for $\gamma_0\theta$ in Z_3, there are at most $f(\hat{n}+j)-j$ possibilities for x. We look at those x which are not a factor of such a $\gamma_0\theta\gamma_0$. Then x is made with a suffix of length $u \in [\![1, \hat{n}+j-1]\!] \subset [\![1, 6\tilde{\varepsilon}m]\!]$ of a word $\gamma_0\theta$ of Z_3 concatenated with a prefix of length $\hat{n}+j-u \in [\![1, 6\tilde{\varepsilon}m]\!]$ of a word $\gamma_0\theta'$ of Z_3 and we nust have $\hat{n}+j-u > \hat{n}$, otherwise x would be a factor of $\gamma_0\theta\gamma_0$. As $Z_3 \subset Z_2$, the suffix is in γ_1 and the prefix in γ_2, so the number of these possible x is at most the number of possible u, which range between 1 and j. Thus the total number of different x is at most $f(\hat{n}+j)-j+j$. □

We shrink our set again.

Let $n \in [\![6\tilde{\varepsilon}m, m]\!]$. In average a factor of length n of a word in Z_3 occurs in at most $\frac{m|Z_3|}{f(n)}$ elements of Z_3 (we assume as above that there are $f(n)$ possible words of size n, possibly by adding ghost factors). We consider the $\frac{f(n)}{mn^2}$ factors of length n which occur the least often. In total, these factors occur in at most $\frac{m|Z_3|}{f(n)}\frac{f(n)}{mn^2} = \frac{|Z_3|}{n^2}$ elements of Z_3. We remove these words from Z_3, for any $m \geq n > 6\tilde{\varepsilon}m$, obtaining a set Z_4. We have removed a proportion at most $1/n^2$ of Z_3 for each n with $m \geq n > 6\tilde{\varepsilon}m \geq 3\hat{n}$, thus a total

[4] For any $x \in]0, +\infty[$, we have $\frac{1}{1-\exp(-x)} < 1+\frac{1}{x}$.

proportion at most $\frac{1}{3\hat{n}} \leq \frac{\delta}{3} < \frac{1}{630}$ of Z_3. This is smaller than $1 - \exp(-\tilde{\varepsilon}hm)$ by Lemma 4.11, thus

$$|Z_4| \geq \exp((1 - 15\tilde{\varepsilon})hm).$$

We can now control medium length factors, using again the missing factors we have just created and the words γ_1 and γ_2 (but not γ_0).

Claim 4.15. *We have $p_{w(Z_4)}(n) \leq f(n)$ for any $n \in [\![1, m]\!]$.*

Proof. Claim 4.14 is still valid for $Z_4 \subset Z_3$. Let $6\tilde{\varepsilon}m \leq n \leq m$ and x a factor of length n of $w(Z_4)$. If x is a factor of a word in Z_4, by construction of Z_4 the number of different possible x is at most $f(n) - \frac{f(n)}{mn^2}$.

If x is not a factor of a word in Z_4, then it is made with a suffix of length $u \in [\![1, n-1]\!]$ of a word of Z_4 concatenated with a prefix of length $n - u$ of another word of Z_4. If $u \geq 6\tilde{\varepsilon}m$, let $u_1 = u - 6\tilde{\varepsilon}m$, $u_2 = n - u$. If $u < 6\tilde{\varepsilon}m$, let $u_1 = 0$, $u_2 = n - 6\tilde{\varepsilon}m$. Our x is made with a variable word of length u_1, concatenated with a factor of $\gamma_1\gamma_2$ which depends only on u, $1 \leq u \leq n-1$, concatenated with a variable word of length u_2. The u_i depend also only on u and, for u_i fixed, the number of possible words of length u_i is at most q_{u_i}. By Lemma 4.7, we get

$$q_{u_1} q_{u_2} \leq \exp(h(2n_r + u_1 + u_2)) \leq \exp(h(2n_r + n - 6\tilde{\varepsilon}m)).$$

Thus by Lemma 4.5 the number of possible x which are not factors of words in Z_4 is at most

$$\begin{aligned} n \exp(h(2n_r + n - 6\tilde{\varepsilon}m)) &\leq m \exp(h(2n_r + n - 6\tilde{\varepsilon}m)) \\ &\leq m \exp(2n_r h - 6h\tilde{\varepsilon}m) \exp(E_0(1 + \frac{\delta}{2})n) \\ &< m \exp(2n_r h + (E_0 \frac{\delta}{2} - 6h\tilde{\varepsilon})m) \exp(E_0 n). \end{aligned}$$

As $n_r < \delta m$, this number is strictly smaller than

$$m \exp(2hm\delta + (E_0 \frac{\delta}{2} - 6h\tilde{\varepsilon})m) \exp(E_0 n) < m \exp(-4h\tilde{\varepsilon}m) \exp(E_0 n)$$

because $\delta \frac{E_0}{2} < h\tilde{\varepsilon}$ by Lemma 4.5 and $\delta < \frac{\tilde{\varepsilon}}{14}$ by (8). By Lemma 4.11, our last estimate on the number of possible x is at most

$$\frac{\exp(E_0 n)}{m^3} \leq \frac{f(n)}{m^3} \leq \frac{f(n)}{mn^2}$$

and our claim is proved. □

Finally we put $Z_5 = Z_4$ if $|Z_4| \leq \exp((1-4\tilde{\varepsilon})hm)$, otherwise we take for Z_5 any subset of Z_4 with $\lceil \exp((1-4\tilde{\varepsilon})hm) \rceil$ elements. In both cases we have

$$|Z_5| \geq \exp((1-\varepsilon)hm).$$

For the long factors, we use mainly the fact that there are many missing factors of length m, but we need also some help from γ_1 and γ_2

Claim 4.16. *We have $p_{w(Z_5)}(n) \leq f(n)$ for any n.*

Proof. Claim 4.15 is still valid for $Z_5 \subset Z_4$. Let x be a factor of $w(Z_5)$ of length $n > m$, with $n = Qm + u$, $0 \leq u < m$, $Q \geq 1$ and thus

$$Qm \geq n/2.$$

The word x is made with a suffix of length u_1 of a word of Z_5, concatenated with Q' words of Z_5 concatenated with a prefix of length u_2 of a word of Z_5. According to the value of u_1, there are two possibilities:

- first case $Q' = Q$ and $u_1 + u_2 = u$ and this occurs for m_1 possible values of u_1;
- second case $Q' = Q - 1$ and $u_1 + u_2 = m + u$ and this occurs for $m - m_1$ values of u_1.

In the first case we bound $q_{u_1} q_{u_2}$ by Lemma 4.7 and the number of possible x by

$$p_1 = m_1 \exp(hu) \exp((1-4\tilde{\varepsilon})hmQ + 2hn_r).$$

Thus $p_1 = m_1 \exp(hn) \exp(-4\tilde{\varepsilon}hmQ + 2hn_r)$, where n_r is at most δm and Qm is at least $n/2$, thus

$$p_1 \leq m_1 \exp(hn) \exp(2hm\delta) \exp(-2\tilde{\varepsilon}hn).$$

In the second case, either the initial suffix of length u_1 or the final prefix of length u_2 contains one of the fixed words γ_1 or γ_2 of length $6\tilde{\varepsilon}m$ and, using again Lemma 4.7, we bound the number of possible x by

$$p_2 = (m - m_1) \exp(h(m+u) - 6h\tilde{\varepsilon}m) \exp((1-4\tilde{\varepsilon})hm(Q-1) + 2hn_r).$$

We have $h(m+u) + hm(Q-1) = hn$ and use $n_r < \delta m$, thus p_2 is at most

$$(m - m_1) \exp((-6\tilde{\varepsilon} + 2\delta)hm) \exp(hn) \exp(-4\tilde{\varepsilon}h(Q-1)m)$$

$$\leq (m - m_1)\exp((-2\tilde{\varepsilon} + 2\delta)hm)\exp(hn)\exp(-4\tilde{\varepsilon}hQm)$$

$$\leq (m - m_1)\exp((-2\tilde{\varepsilon} + 2\delta)hm)\exp(hn)\exp(-2\tilde{\varepsilon}hn)$$

$$\leq (m - m_1)\exp((2\delta hm)\exp(hn)\exp(-2\tilde{\varepsilon}hn).$$

Finally, we have

$$p_n(w(Z_5)) \leq m\exp(hn)\exp(2hm\delta)\exp(-2\tilde{\varepsilon}hn).$$

By Lemma 4.11 and Lemma 4.5 we have $m \leq \exp(\tilde{\varepsilon}hm) \leq \exp(\tilde{\varepsilon}hn)$ (because $m \leq n$). Thus

$$p_n(w(Z_5)) \leq \exp(E_0 n)\exp(2hm\delta)\exp(-\tilde{\varepsilon}hn)\exp(nE_0\frac{\delta}{2})$$

$$\leq \exp(E_0 n)\exp(h(2m+n)\delta)\exp(-\tilde{\varepsilon}hn)$$

$$\leq \exp(E_0 n)\exp(3hn\delta)\exp(-\tilde{\varepsilon}hn).$$

As we have $\tilde{\varepsilon} > 3\delta$ by (8) we get $p_n(w(Z_5)) \leq \exp(E_0 n) \leq f(n)$. □

In view of the considerations at the beginning of the proof of Proposition 4.6, Claim 4.16 completes the proof of Proposition 4.6 and thus of Theorem 4.1. □

References

[Ch33] D. G. Champernowne, The construction of decimals normal in the scale of ten, *J. London Math. Soc.* 8 (1933), 254 260.

[Fal90] K. J. Falconer, Fractal geometry. mathematical foundations and applications John Wiley & Sons, Chichester 1990.

[Fek23] M. Fekete, Uber der Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, *Mathematische Zeitschrift* 17 (1923), 228-249.

[Fer99] S. Ferenczi, Complexity of sequences and dynamical systems, *Discrete Math.*, 206(1-3):145-154, 1999.

[Gri73] C. Grillenberger, Construction of strictly ergodic systems I. Given entropy, *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 25: 323-334, 1973.

[MM10] C. Mauduit and C. G. Moreira, Complexity of infinite sequences with zero entropy, *Acta Arithmetica* 142 (2010), 331-346.

[MM12] C. Mauduit and C. G. Moreira, Generalized Hausdorff dimensions of sets of real numbers with zero entropy expansion, *Ergodic Theory and Dynamical Systems* 32 (2012), 1073-1089.

[MM17] C. Mauduit and C. G. Moreira, Complexity and fractal dimensions for infinite sequences with positive entropy, *preprint*, https://arxiv.org/abs/1702.07698.

[MM18] C. Mauduit and C. G. Moreira, Entropy ratio for infinite sequences with positive entropy, *preprint*, https://arxiv.org/abs/1802.10561.

[MS98] C. Mauduit and A. Sárközy, On finite binary pseudorandom sequences. II. The Champernowne, Rudin-Shapiro and Thue-Morse sequence, a further construction. *J. Number Theory* 73 (2) (1998), 256-276.

[PF02] N. Pytheas Fogg. Substitutions in dynamics, arithmetics and combinatorics, *Lecture Notes in Mathematics 1794*, Springer, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.

[Que87] M. Queffélec, Substitution dynamical systems — spectral analysis, *Lecture Notes in Mathematics 1294*, Springer, 1987.

[Wal82] P. Walters, An Introduction to Ergodic Theory, *Graduate Texts in Mathematics 79*, Springer, 1982.