Maximal pattern complexity for Toeplitz words

(Ergodic Theory and Dynamical Systems 26 (2006), 1-14)

Nertila GJINI*, Teturo KAMAE†, TAN Bo‡ and XUE Yu-Mei §

Abstract

The notion of the maximal pattern complexity of words is introduced in [KZ1, KZ2]. In this paper, we obtain an almost exact formula for the maximal pattern complexity $p_\alpha^*(k)$ of Toeplitz words α on an alphabet $\mathbb{A}$ defined by a sequence of coding words $(\eta^{(n)})^\infty \in (\mathbb{A} \cup \{?\})^{\mathbb{N}}$ $(n = 1, 2, \cdots)$ including just one ? in their cycles $\eta^{(n)}$. Using this formula, we characterize pattern Sturmian words (i.e. $p_\alpha^*(k) = 2k$ $(\forall k)$) in this class. Moreover, we give a characterization of simple Toeplitz words in the sense of [KZ2] in term of pattern complexity. In the case where $\eta^{(1)} = \eta^{(2)} = \cdots$, we obtain the value $\lim_{k\to\infty} p_\alpha^*(k)/k$. We construct a Toeplitz word $\alpha \in \mathbb{A}^{\mathbb{N}}$ with $\sharp\mathbb{A} = 2$ such that $p_\alpha^*(k) = 2^k$ $(k = 1, 2, \cdots)$, while Toeplitz words in our sense always have discrete spectra.

1 Basic notions.

Let $\mathbb{A}$ be a finite set of letters such that $\sharp\mathbb{A} \geq 2$, which is called an *alphabet*. Let $\alpha \in \mathbb{A}^{\mathbb{N}}$ $(\mathbb{N} := \{0, 1, 2, \cdots\})$ be an (infinite) word on $\mathbb{A}$. Let k be a positive integer. By a *k-window* τ, we mean a subset of $\mathbb{N}$ with cardinality k denoting $\tau = \{\tau_0, \tau_1, \cdots, \tau_{k-1}\}$ with $\tau_0 < \tau_1 < \cdots < \tau_{k-1}$. We usually, but not always, assume for a window τ that $\tau_0 = 0$. The k-window $\tau = \{0, 1, \cdots, k-1\}$ is called the *k-block window.*

For a k-window $\tau = \{\tau_0, \tau_1, \cdots, \tau_{k-1}\}$ and a word $\alpha = \alpha(0)\alpha(1)\alpha(2)\cdots$, we denote

$$\begin{aligned} \alpha[n+\tau] &:= \alpha(n+\tau_0)\alpha(n+\tau_1)\cdots\alpha(n+\tau_{k-1}) \\ F_\alpha(\tau) &:= \{\alpha[n+\tau]\ ;\ n = 0, 1, 2, \cdots\} \\ p_\alpha(\tau) &:= \sharp F_\alpha(\tau). \end{aligned}$$

*University of New York, Tirana, Albania (ngjini@unyt.edu.al)

†Matsuyama University, 790-8578 Japan (kamae@apost.plala.or.jp)

‡Department of Mathematics, Huazhong University of Science and Technology, Wuhan 430074, P.R. China (bo_tan@163.com)

§Department of Mathematics, Tsinghua University, Beijing 100084, P.R.China (yxue@math.tsinghua.edu.cn)

An element in $F_\alpha(\tau)$ is called a *τ-factor* of α. The *maximal pattern complexity* p^*_α for a word α has been introduced by the second author together with Zamboni [KZ1] as

$$p^*_\alpha(k) := \sup_\tau p_\alpha(\tau) \quad (k = 1, 2, 3, \cdots),$$

where the supremum is taken over all k-windows τ, while the *block complexity* p_α is defined as

$$p_\alpha(k) = p_\alpha(\{0, 1, \cdots, k-1\}).$$

It is known (Morse and Hedlund [MH]) that for a word α, the following statements are equivalent:
(i) *α is eventually periodic,*
(ii) *$p_\alpha(k)$ is bounded in k,*
(iii) *$p_\alpha(k) < k+1$ for some $k = 1, 2, \cdots$,*
while the following parallel statements with respect to the maximal pattern complexity are equivalent [KZ1]:
(i) *α is eventually periodic,*
(ii') *$p^*_\alpha(k)$ is bounded in k,*
(iii') *$p^*_\alpha(k) < 2k$ for some $k = 1, 2, \cdots$.*

A word α with block complexity $p_\alpha(k) = k+1$ $(k = 1, 2, 3, \cdots)$ is known as a *Sturmian word* and is studied extensively (see for example Berthé [B] and the references therein). A word α with maximal pattern complexity $p^*_\alpha(k) = 2k$ $(k = 1, 2, 3, \cdots)$ is called a *pattern Sturmian word* and is studied in [KZ1, KZ2]. It is known that Sturmian words are always pattern Sturmian words, while simple Toeplitz words defined below are pattern Sturmian words which are not Sturmian words.

The value $\lim_{k\to\infty}(1/k)\log p^*_\alpha(k)$ is known to be an invariant of the topological dynamical system arising from α taking values in $\{\log n;\ n = 1, 2, \cdots, \sharp\mathbb{A}\}$ (see Huang and Ye [HY]). If the measure-theoretic dynamical system arising from α has a partially continuous spectrum, then this value is nonzero, but the converse is not true.

For a nonempty set A of letters and a letter ? which is not included in A, let $\mathcal{P}(A, ?)$ be the set of periodic words $\xi = \eta^\infty \in (A \cup \{?\})^{\mathbb{N}}$ such that every letter in A occurs at least once in η while ? occurs in η just once. Hence, the length of η is the *minimum period* of ξ. For $\xi^1 \in \mathcal{P}(A, ?)$ and $\xi^2 \in \mathcal{P}(B, ?)$, we define $\xi^1 \lhd \xi^2$ by substituting every occurrence of ? in ξ^1 by $\xi^2(0),\ \xi^2(1),\ \xi^2(2),\ \cdots$ in the order. Then $\xi^1 \lhd \xi^2 \in \mathcal{P}(A \cup B, ?)$.

Let $\xi^n \in \cup_{\emptyset \neq S \subset \mathbb{A}} \mathcal{P}(S, ?)$ $(n = 1, 2, \cdots)$. We define

$$\alpha = \xi^1 \lhd \xi^2 \lhd \xi^3 \lhd \cdots \in (\mathbb{A} \cup \{?\})^{\mathbb{N}}$$

as the limit of $\xi^1 \lhd \xi^2 \cdots \lhd \xi^n$ as n tends to ∞. If the first occurrence place of ? in ξ^n is not 0 for infinitely many n's, then $\alpha \in \mathbb{A}^{\mathbb{N}}$. In this case, we call α a *Toeplitz* *word with single hole* provided that it is not eventually periodic, and ξ^1, ξ^2, $\cdots$ a sequence of its *coding words*. Here, we only consider Toeplitz words with single hole (that is, one ? in the minimum cycles of all the coding words), so that we call them simply *Toeplitz words*. Since any Toeplitz word α is not eventually periodic, $p^*_\alpha(k) \geq 2k$ $(k = 1, 2, \cdots)$ holds by the above equivalence.

We remark that any of $\alpha = \xi^1 \lhd \xi^2 \lhd \xi^3 \lhd \cdots \in \mathbb{A}^{\mathbb{N}}$ with $\xi^n \in \cup_{\emptyset \neq S \subset \mathbb{A}} \mathcal{P}(S, ?)$ $(n = 1, 2, \cdots)$ is recurrent. Moreover, it is periodic if and only if there exist $a \in \mathbb{A}$ and $h \geq 1$ such that $\xi^i \in \mathcal{P}(\{a\}, ?)$ for any $i \geq h$.

To prove "if" part, let $\xi^i \in \mathcal{P}(\{a\}, ?)$ for any $i \geq h$. Then, $\xi^h \lhd \xi^{h+1} \lhd \cdots = aaa\cdots$ holds, so that α is obtained from the periodic word $\xi^1 \lhd \xi^2 \lhd \cdots \lhd \xi^{h-1}$ by filling all the ? by the same letter a. Hence, α is periodic.

Conversely, assume that $\sharp S \geq 2$, where S is the set of $a \in \mathbb{A}$ which appears in ξ^i for infinitely many i's. Then, for any sufficiently large h, $\xi^h \lhd \xi^{h+1} \lhd \cdots$ consists only of letters in S, any of which appears infinitely often. On the other hand, the minimum period L of the periodic word $\zeta := \xi^1 \lhd \xi^2 \lhd \cdots \lhd \xi^{h-1}$ containing just one ? in its minimum cycle tends to ∞ as h tends to ∞. In α, all the ? in ζ are filled by the letters in S, while any letter in S is used infinitely often. This shows first that α is recurrent. Also, this shows that we can find 2 blocks in α with an arbitrary length $K \leq L$ located at the same place modulo L having different numbers of a's (and b's) if $a, b \in S$ with $a \neq b$, which implies that K is not a period of α. Since we can take $K = 1, 2, \cdots$, α is not periodic.

A Toeplitz word $\alpha \in \mathbb{A}^{\mathbb{N}}$ is called a *simple* Toeplitz word if $\sharp\mathbb{A} = 2$, say $\mathbb{A} = \{a, b\}$, and it has a sequence of coding words $\xi_1, \xi_2, \cdots \in \mathcal{P}(\{a\}, ?) \cup \mathcal{P}(\{b\}, ?)$. It is known that if α is a simple Toeplitz word, then it is a pattern Sturmian word, that is $p^*_\alpha(k) = 2k$ $(k = 1, 2, \cdots)$. We remark that the proof for it in [KZ2] is wrong. Here, we give not only a correct proof for it but also a characterization of the pattern Sturmian words among all Toeplitz words. We also give a characterization for a word to be a simple Toeplitz word in term of pattern complexity.

2 Main results

Let $\beta \in \mathbb{A}^{\mathbb{N}}$ be a recurrent word and $\xi \in \mathcal{P}(\mathbb{A}, ?)$ with the minimum period r. Let $\tau = \{0 = \tau_0 < \tau_1 < \cdots < \tau_{k-1}\}$ be a k-window. We decompose τ into a union of subwindows

$$\tau = \bigcup_{i \in L} \tau^i, \tag{2.1}$$

where for $i = 0, 1, \cdots, r-1$,

$$\tau^i := \{\tau_j \ ; \ j = 0, 1, \cdots, k-1 \text{ such that } \tau_j \equiv i \pmod{r}\} \text{ and}$$
$$L := \{i \in \{0, 1, \cdots, r-1\} \ ; \ \tau^i \neq \emptyset\}.$$

Here, we also denote

$$\overline{\tau}^i := (\tau^i - i)/r = \{(\tau_j - i)/r \ ; \ \tau_j \in \tau^i\} \quad (i \in L). \tag{2.2}$$

For $a \in \mathbb{A} \cup \{?\}$, $S \subset \mathbb{A}$, $\eta_1 \cdots \eta_n \in (\mathbb{A} \cup \{?\})^n$ and $U \subset (\mathbb{A} \cup \{?\})^n$ with $n \in \mathbb{N}$, we denote

$$\begin{aligned} \pi_S(a) &= \begin{cases} \{a\} & \text{if } a \neq ? \\ S & \text{if } a = ? \end{cases} \\ \pi_S(\eta_1 \cdots \eta_n) &= \{\zeta_1 \cdots \zeta_n \ ; \ \zeta_i \in \pi_S(\eta_i) \ (i = 1, \cdots, n)\} \\ \pi_S U &= \bigcup_{u \in U} \pi_S(u). \end{aligned}$$

We define

$$\begin{aligned} D(\xi, L) &:= \sharp \pi_{\mathbb{A}} F_\xi(L) - \ell \sharp \mathbb{A} \\ E(\xi, L) &:= \sharp \left(\pi_{\mathbb{A}} F_\xi(L) \cup \{a^\ell ; \ a \in \mathbb{A}\} \right) - \ell \sharp \mathbb{A}, \end{aligned} \tag{2.3}$$

where we denote $\ell = \sharp L$ and $F_\xi(L)$ is the set of L-factors of ξ.

Let $\alpha = \xi \lhd \beta \in \mathbb{A}^{\mathbb{N}}$. Let i_0 be such that $0 \leq i_0 \leq r - 1$ and $\xi_{i_0} = ?$. Then, $\xi_i = ?$ if and only if $i \equiv i_0 \pmod r$. By a *constant* word, we mean a word in $\{a^n ; \ a \in \mathbb{A}\}$ for some $n \in \mathbb{N}$. For $U \subset \mathbb{A}^n$ with $n \in \mathbb{N}$, we denote by $C(U)$ the set of letters appearing in the constant words in U. It is easy to prove the following Lemma.

Lemma 1. *Let $\alpha = \xi \lhd \beta$, where $\xi \in \mathcal{P}(\mathbb{A}, ?)$ with the minimum period r and $\beta \in \mathbb{A}^{\mathbb{N}}$ is a recurrent word. Let τ be a k-window with the decomposition* (2.1) *and* (2.2)*. Then for any $i \in L$, we have*

$$\begin{gathered} F_\beta(\overline{\tau}^i) \subset F_\alpha(\tau^i) \\ C(F_\beta(\overline{\tau}^i)) \subset C(F_\alpha(\tau^i)) = \mathbb{A} \\ \sharp \left(F_\alpha(\tau^i) \setminus F_\beta(\overline{\tau}^i) \right) = \sharp \mathbb{A} - \sharp C(F_\beta(\overline{\tau}^i)). \end{gathered}$$

We denote

$$\begin{aligned} F &:= \{u \in F_\alpha(\tau) \ ; \ u|_{\tau^i} \text{ is a constant word for any } i \in L\}, \\ F^i &:= \{u \in F_\alpha(\tau) \ ; \ u|_{\tau^i} \text{ is not a constant word}\} \quad (i \in L), \end{aligned}$$

where for $u \in F_\alpha(\tau)$ with $u = \alpha[n + \tau]$, we denote $u|_{\tau^i} = \alpha[n + \tau^i]$. If $u \in F^i$ with $i \in L$, then $u = \alpha[n + \tau]$ holds with n such that $n + i \equiv i_0 \pmod r$ since otherwise $u|_{\tau^i}$ is a constant word. Moreover, if $u \in F^i$ and $i' \in L$ with $i \neq i'$, then $u|_{\tau^{i'}}$ is a fixed constant word with letter $\xi(j)$ such that $j \equiv i_0 + i' - i \pmod r$ which is independent of $u \in F^i$. Hence, F and F^i's are disjoint each other and any of $u \in F^i$ is determined by $u|_{\tau^i}$.

Note that if $u \in F^i$ with $u = \alpha[n + \tau]$ and $n + i - i_0 \geq 0$, then

$$u|_{\tau^i} = \beta[(n + i - i_0)/r + \overline{\tau}^i].$$

Conversely, if $v \in F_\beta(\overline{\tau}^i)$ is not a constant word with $v = \beta[n + \overline{\tau}^i]$, then $v = u|_{\tau^i}$ holds for $u \in F^i$ with $u = \alpha[rn + i_0 - i + \tau^i]$. Thus, we have

$$\sharp F^i = p_\beta(\overline{\tau}^i) - \sharp C(F_\beta(\overline{\tau}^i)) = p_\alpha(\tau^i) - \sharp \mathbb{A},$$

where the second equality holds by Lemma 1.

Consider the set F. To any word $u \in F$, we associate a new word $\tilde{u} \in F_\alpha(L)$ with the property $\tilde{u}(i) \in C(u|_{\tau^i})$ for all $i \in L$. Denote

$$\tilde{F} = \{\tilde{u} \in F_\alpha(L);\ u \in F\},$$

then, by the definition of F, it is easy to see that there is a bijection between the sets F and $\tilde{F}$. Also it holds that

$$\tilde{F} = (F_\xi(\tau) \cap \mathbb{A}^l) \cup \bigcup_{i \in L} \pi_{C_i}(\xi[i_0 - i + L]),$$

with $C_i := C(F_\beta(\overline{\tau}^i))$. Moreover, it holds that $C(F_\alpha(\tau)) \subset C(\pi_{\mathbb{A}}(F_\xi(\tau))$ and

$$\{a^l;\ a \in C(\pi_{\mathbb{A}}(F_\xi(\tau)))\} \setminus \{a^l;\ a \in C(F_\alpha(\tau))\} \subset \pi_{\mathbb{A}}(F_\xi(L)) \setminus \tilde{F}.$$

Therefore, we have

$$\tilde{F} \subset \pi_{\mathbb{A}}(F_\xi(L)) \setminus \left(\{a^l;\ a \in C(\pi_{\mathbb{A}}(F_\xi(\tau)))\} \setminus \{a^l;\ a \in C(F_\alpha(\tau))\}\right),$$

while $\tilde{F} = \pi_{\mathbb{A}}(F_\xi(L))$ holds if $C_i = \mathbb{A}$ for any $i \in L$. Hence, we have

$$\begin{aligned}
p_\alpha(\tau) &= \sharp F + \sum_{i \in L} \sharp F^i = \sharp \tilde{F} + \sum_{i \in L} \sharp F^i \\
&\leq \sharp \pi_{\mathbb{A}} F_\xi(\tau) - \sharp C(\pi_{\mathbb{A}} F_\xi(\tau)) + \sharp C(F_\alpha(\tau)) + \sum_{i \in L} (p_\alpha(\tau^i) - \sharp \mathbb{A}) \\
&= \sharp \pi_{\mathbb{A}} F_\xi(L) - \sharp C(\pi_{\mathbb{A}} F_\xi(L)) + \sharp C(F_\alpha(\tau)) + \sum_{i \in L} (p_\alpha(\tau^i) - \sharp \mathbb{A}) \\
&= D(\xi, L) + \sum_{i \in L} p_\alpha(\tau^i) - \sharp C(\pi_{\mathbb{A}} F_\xi(L)) + \sharp C(F_\alpha(\tau)),
\end{aligned}$$

while $p_\alpha(\tau) = D(\xi, L) + \sum_{i \in L} p_\alpha(\tau^i)$ holds if $C_i = \mathbb{A}$ for any $i \in L$.

Thus, we get the following Theorem 1.

Theorem 1. *For any $\xi \in \mathcal{P}(\mathbb{A}, ?)$ with the minimal period r and a recurrent word $\beta \in \mathbb{A}^{\mathbb{N}}$, let $\alpha = \xi \lhd \beta$ and τ be a k-window with the decomposition in* (2.1) *and* (2.2)*. Then, we have*

$$p_\alpha(\tau) - \sharp C(F_\alpha(\tau)) \leq D(\xi, L) - \sharp C(\pi_{\mathbb{A}}(F_\xi(L))) + \sum_{i \in L} p_\alpha(\tau^i).$$

Particularly, we have

$$p_\alpha(\tau) \leq D(\xi, L) + \sum_{i \in L} p_\alpha(\tau^i)$$

with the equality if $C(F_\beta(\overline{\tau}^i)) = \mathbb{A}$ for any $i \in L$.

Corollary 1. *With the same setting as in* Theorem 1*, we have the following statements.*
(1) *It holds that* $C(F_\alpha(\tau^i)) = \mathbb{A}$ *and* $p_\alpha(\tau^i) = p_\beta(\overline{\tau}^i) + \sharp\mathbb{A} - \sharp C(F_\beta(\overline{\tau}^i))$ *for any* $i \in L$*. In particular, we have* $p_\alpha(r\tau) = p_\beta(\tau) + \sharp\mathbb{A} - \sharp C(F_\beta(\tau))$*.*
(2) $p^*_\alpha(k) \geq p^*_\beta(k)$ *holds for any* $k = 1, 2, \cdots$ *.*
(3) *If* $\beta = \eta \lhd \gamma$ *holds with* $\eta \in \mathcal{P}(\mathbb{A}, ?)$ *having the minimum period* s *and a recurrent word* $\gamma \in \mathbb{A}^{\mathbb{N}}$ *satisfying the following* Condition$(*)$ *for* τ*, then we have the equality*

$$p_\alpha(\tau) = D(\xi, L) + \sum_{i \in L} p_\alpha(\tau^i) = D(\xi, L) + \sum_{i \in L} p_\beta(\overline{\tau}^i).$$

Condition$(*)$ *For any* $i \in L$ *and* $j, j' \in \tau^i$*,* rs *divides* $j - j'$*.*

Proof. The first part of (1) follows from Lemma 1. The second part follows from the first part as $(r\tau)^0 = r\tau$ and $\overline{(r\tau)}^0 = \tau$.

For an arbitrary k-window η, put $\tau = r\eta$. Then, since $\tau^0 = \tau$ and $\overline{\tau}^0 = \eta$, we have $p_\alpha(\tau) \geq p_\beta(\eta)$ by (1), which proves (2).

Let us prove (3). Since $\overline{\tau}^i$ for any $i \in L$ is contained in single modulo class of s by the Condition$(*)$, $F_\beta(\overline{\tau}^i)$ contains all constant words with letters appearing in η, that is $\mathbb{A}$ as $\eta \in \mathcal{P}(\mathbb{A}, ?)$. Hence, $C(F_\beta(\overline{\tau}^i)) = \mathbb{A}$ for any $i \in L$, and by Theorem 1 and (1), we have the equality

$$\begin{aligned} p_\alpha(\tau) &= D(\xi, L) + \sum_{i \in L} p_\alpha(\tau^i) \\ &= D(\xi, L) + \sum_{i \in L} p_\beta(\overline{\tau}^i). \end{aligned}$$

□

Lemma 2. *Let* $\alpha \in \mathbb{A}^{\mathbb{N}}$ *be any simple Toeplitz word. Then,*

$$p_\alpha(\tau) - \sharp C(F_\alpha(\tau)) + \#\mathbb{A} \leq 2k \tag{2.4}$$

holds for any k*-window* τ*. In particular, a simple Toeplitz word is a pattern Sturmian word.*

Proof. To prove (2.4) for any simple Toeplitz word $\alpha \in \mathbb{A}^{\mathbb{N}}$ with $\mathbb{A} = \{a, b\}$ and any k-window, we use the induction on k.

If $k = 1$ or 2, (2.4) holds clearly.

For $k \geq 3$, assume that (2.4) holds for any ℓ-window with $\ell < k$ and any simple Toeplitz word. Take a simple Toeplitz word $\alpha \in \mathbb{A}^{\mathbb{N}}$ and a k-window τ. Since α is a simple Toeplitz word, we may assume without loss of generality that $\alpha = \xi \lhd \beta$ holds for a simple Toeplitz word $\beta \in \mathbb{A}^{\mathbb{N}}$ and $\xi \in \mathcal{P}(\mathbb{A}, ?)$ such that $\xi = \eta \lhd \zeta$ with $\eta \in \mathcal{P}(\{a\}, ?)$ and $\zeta \in \mathcal{P}(\{b\}, ?)$. Let the minimum periods of η and ζ be s and t, respectively. Thus the minimum period of ξ is $r = st$.

Case 1 If τ is not a multiple of r, then we have the decomposition of τ as in (2.1) and (2.2) with $l = \sharp L \geq 2$. By Theorem 1, Lemma 1 and the assumption of the induction, we have

$$\begin{aligned} p_\alpha(\tau) - C + 2 &\leq D(\xi, L) - C' + 2 + \sum_{i\in L} p_\alpha(\tau^i) \\ &= D(\xi, L) - C' + 2 + \sum_{i\in L} \left(p_\beta(\overline{\tau}^i) + 2 - \sharp C_i\right) \\ &\leq D(\xi, L) - C' + 2 + \sum_{i\in L} 2\sharp\overline{\tau}^i \\ &= D(\xi, L) - C' + 2 + 2k, \end{aligned}$$

where we put $C := \sharp C(F_\alpha(\tau))$ and $C' := \sharp C(\pi_{\mathbb{A}} F_\xi(L))$. Therefore, to prove (2.4), it is sufficient to prove

$$D(\xi, L) - C' + 2 \leq 0. \tag{2.5}$$

Decompose $L = \{L_1, L_2, \cdots, L_\ell\}$ into the modulo classes of s as follows.

$$L = \bigcup_{i\in K} L^i,$$

where for $i = 0, 1, \cdots, s-1$.

$$\begin{aligned} L^i &:= \{L_j \ ; \ j = 0, 1, \cdots, k-1 \text{ such that } L_j \equiv i \ (\mathrm{mod}\ s)\} \text{ and} \\ K &:= \{i = 0, 1, \cdots, s-1 \ ; \ L^i \neq \emptyset\}. \end{aligned}$$

For $i \in K$, define

$$G^i = \{u \in F_\xi(L) \ ; \ u = \xi[n + L] \text{ with } n \text{ such that } n \equiv i_0 - i \ (\mathrm{mod}\ s)\},$$

where i_0 is such that $0 \leq i_0 \leq s-1$ and $\eta(i_0) = ?$. We also define

$$G = \{u \in F_\xi(L) \ ; \ u = \xi[n + L] \text{ with } n \text{ such that } n \notin i_0 - K \ (\mathrm{mod}\ s)\}.$$

Note that $G = \{a^\ell\}$ if $G \neq \emptyset$. If $u \in G^i$ with $i \in K$, then $u|_{L\setminus L^i} = a^{\ell - \sharp L_i}$ holds, while $u|_{L^i}$ consists only of b's except possibly for at one place. Therefore, for any $i \in K$, we have $\sharp \pi_{\mathbb{A}} G_i = \sharp L^i + 1$. Moreover, $\pi_{\mathbb{A}} G_i$ contains the constant word a^ℓ if and only if $\sharp L^i = 1$. Hence, we have

$$\sharp\left(\pi_{\mathbb{A}} G^i \setminus \{a^\ell\}\right) = \sharp L^i + 1_{\sharp L^i \geq 2}.$$

Define $1_a = 1$ or $= 0$ according to whether $\pi_{\mathbb{A}} F_\xi(L)$ contains the constant word a^ℓ or not. In the same way, we define $1_b = 1$ or $= 0$ according to whether $\pi_{\mathbb{A}} F_\xi(L)$

contains the constant word b^ℓ or not. Then, we have $C' = 1_a + 1_b$. On the other hand, we have

$$\begin{aligned}\sharp \pi_A F_\xi(L) &= \sharp\left(G \cup \bigcup_{i\in K} \pi_{\mathbb{A}} G^i\right)\\ &\leq 1_a + \sum_{i\in K}(\sharp L_i + 1_{\sharp L^i \geq 2})\\ &= 1_a + \ell + \sum_{i\in K} 1_{\sharp L^i\geq 2}\\ &\leq 1_a + \ell + \lfloor \ell/2 \rfloor.\end{aligned}$$

Thus, we have

$$D(\xi, L) - C' + 2 \leq 2 - 1_b - \ell + \sum_{i\in K} 1_{\sharp L^i \geq 2} \leq 2 - 1_b - \lceil \ell/2 \rceil. \tag{2.6}$$

If $\ell \geq 3$, then (2.5) follows from (2.6). If $\ell = 2$ with $\sharp K = 2$, then since $\sum_{i\in K} 1_{\sharp L^i\geq 2} = 0$, we have (2.5) by (2.6). Finally, if $\ell = 2$ with $\sharp K = 1$, then we have 2 cases, either $\pi_{\mathbb{A}} F_\xi(L) = \{aa, ab, ba\}$ or $\{aa, ab, ba, bb\}$. In the first case, we have $C' = 1$, while in the second case, we have $C' = 2$. Any case, we have $\sharp \pi_{\mathbb{A}} F_\xi(L) - C' = 2$. Thus, $D(\xi, L) - C' + 2 = 0$, which completes the proof of Case 1.

Case 2 If τ is a multiple of r, then by (1) of Corollary 1, we have

$$p_\alpha(\tau) - \sharp C(F_\alpha(\tau)) + 2 = p_\beta(\tau/r) - \sharp C(F_\beta(\tau/r)) + 2.$$

If τ is a multiple of r^e but not a multiple of r^{e+1}, then we repeat this argument e times until we get a simple Toeplitz word γ such that

$$p_\alpha(\tau) - \sharp C(F_\alpha(\tau)) + 2 = p_\gamma(\tau/r^e) - \sharp C(F_\gamma(\tau/r^e)) + 2.$$

Since τ/r^e is not a multiple of r, we get

$$p_\alpha(\tau) - \sharp C(F_\alpha(\tau)) + 2 = p_\gamma(\tau/r^e) - \sharp C(F_\gamma(\tau/r^e)) + 2 \leq 2k$$

by Case 1, which completes the first part of Theorem 2.

Since $-\sharp C(F_\alpha(\tau)) + \#\mathbb{A} \geq 0$, (2.4) for any simple Toeplitz word α and any k-window τ implies that $p^*_\alpha(k) \leq 2k$ $(k = 1, 2, \cdots)$ for any simple Toeplitz word α. Here, the equality follows since any Toeplitz word is not eventually periodic. Thus, any simple Toeplitz word is a pattern Sturmian word. □

Theorem 2. *Let $\alpha \in \mathbb{A}^{\mathbb{N}}$ be a Toeplitz word. Then, it is a pattern Sturmian word if and only if either it is a simple Toeplitz word or there exists a simple Toeplitz word $\gamma \in \mathbb{A}^{\mathbb{N}}$ and $\xi \in \mathcal{P}(\mathbb{A}, ?)$ with the minimum period r such that $D(\xi, L) \leq 0$ for any $L \subset \{0, 1, \cdots, r-1\}$, and $\alpha = \xi \lhd \gamma$.*

Proof. We prove the "if" part first. By Lemma 2, it is sufficient to prove that for a simple Toeplitz word $\gamma \in \mathbb{A}^{\mathbb{N}}$, and $\xi \in \mathcal{P}(\mathbb{A}, ?)$ with the minimum period r such that $D(\xi, L) \leq 0$ for any $L \subset \{0, 1, \cdots, r-1\}$, $\alpha := \xi \lhd \gamma$ is a pattern Sturmian word.

Take any k-window τ and use the notations in (2.1) and (2.2). We have already proved in Lemma 2 that

$$p_\gamma(\overline{\tau}^i) - \sharp C(F_\gamma(\overline{\tau}^i)) + 2 \leq 2\#\overline{\tau}^i.$$

Then, by Theorem 1 and (1) of Corollary 1, we have

$$\begin{aligned} p_\alpha(\tau) &\leq D(\xi, L) + \sum_{i \in L} p_\alpha(\tau^i) \\ &= D(\xi, L) + \sum_{i \in L} \left(p_\beta(\overline{\tau}^i) - \sharp C(F_\gamma(\overline{\tau}^i)) + 2\right) \\ &\leq \sum_{i \in L} 2k_i = 2k, \end{aligned}$$

where we put $k_i = \sharp \tau^i$. Thus, α is a pattern Sturmian word.

Let us prove the "only if" part. Assume that a Toeplitz word α is a pattern Sturmian word. Then, just 2 letters appear in α, so that we may assume that $\mathbb{A} = \{a, b\}$. By the aperiodicity, we may assume that there exists a sequence $\xi^1, \xi^2, \cdots$ of coding words of α consisting only of elements in $\mathcal{P}(\mathbb{A}, ?)$. Let r_i be the minimum period of ξ^i $(i = 1, 2, \cdots)$. Denote

$$\alpha^i = \xi^i \lhd \xi^{i+1} \lhd \cdots \quad (i = 1, 2, \cdots).$$

Then $\alpha^1 = \alpha$ and $\alpha^i = \xi^i \lhd \alpha^{i+1}$, and by (2) of Corollary 1, all α^i's are pattern Sturmian words.

Suppose that $D(\xi^i, L) > 0$ holds for some $i = 1, 2, \cdots$ and $L \subset \{0, 1, \cdots, r_i - 1\}$. Then, by (3) of Corollary 1 with $\tau = L$ as Condition $(*)$ is clearly satisfied, we have

$$p_{\alpha^i}(L) = D(\xi, L) + \sum_{j \in L} p_{\alpha^i}(\{j\}) = D(\xi, L) + 2\ell > 2\ell,$$

where $\ell := \sharp L$. This implies that

$$p_\alpha(r_1 \cdots r_{i-1} L) \geq p_{\alpha^i}(L) > 2\ell$$

by (1) of Corollary 1, which contradicts the assumption that α is a pattern Sturmian word. Thus, we have the following Lemma 3.

Lemma 3. *Let $\xi \in \mathcal{P}(\mathbb{A}, ?)$ be any word appearing in some sequence of coding words consisting only of elements in $\mathcal{P}(\mathbb{A}, ?)$ of a Toeplitz word in $\mathbb{A}^{\mathbb{N}}$ which is a pattern Sturmian word as well. Then, we have*

$$D(\xi, L) \leq 0 \text{ for any } L \subset \{0, 1, \cdots, r-1\}, \tag{2.7}$$

where r is the minimum period of ξ.

Let us return to the proof of the "only if" part of Theorem 2. We call $\xi \in \mathcal{P}(\mathbb{A}, ?)$ a *simple* coding word, if there exist $\eta^i \in \mathcal{P}(\{a\}, ?) \cup \mathcal{P}(\{b\}, ?)$ $(i = 1, 2, \cdots, h)$ such that $\xi = \eta^1 \lhd \eta^2 \lhd \cdots \lhd \eta^h$. Any element in $\mathcal{P}(\{a\}, ?) \cup \mathcal{P}(\{b\}, ?)$ is also called a *simple* coding word. We call $\xi \in \mathcal{P}(\mathbb{A}, ?)$ *irreducible* if there does not exist a simple coding word ζ such that either $\xi = \zeta$ or $\xi = \eta \lhd \zeta$ holds with some $\eta \in \mathcal{P}(\mathbb{A}, ?) \cup \mathcal{P}(\{a\}, ?) \cup \mathcal{P}(\{b\}, ?)$.

Suppose that one of $\xi^2, \xi^3, \cdots$, say ξ^h, is not simple. Then, either ξ^h is irreducible or there exists a simple coding word ζ and an irreducible η such that $\xi^h = \eta \lhd \zeta$. Any case, we have the decomposition that

$$\alpha = \xi \lhd \eta \lhd \beta$$

with $\xi \in \mathcal{P}(\mathbb{A}, ?)$, an irreducible $\eta \in \mathcal{P}(\mathbb{A}, ?)$ and a Toeplitz word β. Hence, by Lemma 3, we have $D(\xi \lhd \eta, rL) \leq 0$ for any $L \subset \{0, 1, \cdots, s-1\}$, where r, s are the minimum periods of ξ and η, respectively.

On the other hand, we will prove the following Lemma 4, which contradicts this fact. This implies that any of $\xi^2, \xi^3, \cdots \in \mathcal{P}(\mathbb{A}, ?)$ in $\alpha = \xi^1 \lhd \xi^2 \lhd \xi^3 \cdots$ is simple. Therefore, $\gamma := \xi^2 \lhd \xi^3 \cdots$ is a simple Toeplitz word. Moreover, we have $\alpha = \xi^1 \lhd \gamma$ with ξ^1 satisfying (2.7) by Lemma 3, which completes the proof of Theorem 2. □

Lemma 4. *For any $\xi, \eta \in \mathcal{P}(\mathbb{A}, ?)$ such that $\sharp\mathbb{A} = 2$ and η is irreducible, we have $D(\xi \lhd \eta, rL) \geq 1$ for $L = \{0, 1, 2, \cdots, s-1\}$, where r, s are the minimum periods of ξ and η, respectively.*

Proof. Let $\mathbb{A} = \{a, b\}$. It is easy to see that

$$\pi_{\mathbb{A}} F_{\xi \lhd \eta}(rL) = \mathbb{A}^s \cup \pi_{\mathbb{A}} F_\eta(L).$$

The set $\pi_{\mathbb{A}} F_\eta(L)$ contains

$$\begin{aligned} &(\eta(n)\eta(n+1)\cdots\eta(n+s-1))_a \quad \text{and} \\ &(\eta(n)\eta(n+1)\cdots\eta(n+s-1))_b \quad (n = 0, 1, \cdots, s-1), \end{aligned}$$

where the word $(\cdots)_a$ or $(\cdots)_b$ denotes the word obtained from the original word $\cdots$ by replacing all occurrences of ? in $\cdots$ by a or b, respectively.

We prove that all $2s$ words in this list are different from each other. Since $(\eta(n)\eta(n+1)\cdots\eta(n+s-1))_a$ and $(\eta(m)\eta(m+1)\cdots\eta(m+s-1))_b$ contain different numbers of letters a for any n, m, they are different from each other. Suppose that

$$(\eta(n)\eta(n+1)\cdots\eta(n+s-1))_a = (\eta(m)\eta(m+1)\cdots\eta(m+s-1))_a$$

holds for some $n < m$. Then with the minimum $s_1 = m - n$ as this, we have the factorization $s = s_1 s_2$ and

$$(\eta(0)\eta(1)\cdots\eta(s-1))_a = (\eta(0)\eta(1)\cdots\eta(s_1-1))_a^{s_2}.$$

There exists just one i_0 with $0 \le i_0 \le s-1$ such that $\eta(i_0) = ?$. Let $i_0 = i_1 + i_2 s_1$ with $0 \le i_1 < s_1$ and $0 \le i_2 < s_2$. Then we can write $\eta = \eta' \triangleleft \eta''$ with

$$\eta'(i) = \begin{cases} \eta(i) & i \not\equiv i_1 \pmod{s_1} \\ ? & i \equiv i_1 \pmod{s_1} \end{cases}$$
$$\eta''(i) = \begin{cases} a & i \not\equiv i_2 \pmod{s_2} \\ ? & i \equiv i_2 \pmod{s_2} \end{cases},$$

which contradicts the assumption that η is irreducible.

Thus, all elements in the list $(\eta(n)\eta(n+1)\cdots\eta(n+s-1))_a$ $(n = 0, 1, \cdots, s-1)$ are different from each other. In the same way, all elements in the list $(\eta(n)\eta(n+1)\cdots\eta(n+s-1))_b$ $(n = 0, 1, \cdots, s-1)$ are different from each other. Therefore, we have $2s$ elements in $\pi_{\mathbb{A}} F_\eta(L)$, which contains at most one constant word.

Thus, we have $\sharp\pi_{\mathbb{A}} F_{\xi\triangleleft\eta}(rL) \ge 2s+1$, which implies that $D(\xi \triangleleft \eta, rL) \ge 1$. ☐

Theorem 3. *Let $\alpha \in \mathbb{A}^{\mathbb{N}}$ be a Toeplitz word. It is a simple Toeplitz word if and only if* (2.4) *holds for any k-window τ.*

Proof. The "only if" part has been already proved in Lemma 2. Here, we prove the "if" part. Assume that (2.4) holds for a Toeplitz word $\alpha \in \mathbb{A}^{\mathbb{N}}$. Then, α is a pattern Sturmian word. Hence, we may assume that $\sharp\mathbb{A} = 2$

Suppose that α is not a simple Toeplitz word. Then by Theorem 2, there exists a simple Toeplitz word $\gamma \in \mathbb{A}^{\mathbb{N}}$ and $\xi \in \mathcal{P}(\mathbb{A}, ?)$ with the minimum period r such that $D(\xi, L) \le 0$ for any $L \subset \{0, 1, \cdots, r-1\}$, and $\alpha = \xi \triangleleft \gamma$. If ξ is a simple coding word, then α is a simple Toeplitz word contradicting our supposition. Hence, ξ is not a simple coding word. Then, there exists an irreducible word ζ such that either $\xi = \zeta$ or there exists a simple coding word ζ' such that $\xi = \zeta \triangleleft \zeta'$. Any case, we have $\alpha = \zeta \triangleleft \beta$ with an irreducible coding word ζ and a simple Toeplitz word β. Let s be the minimum period of ζ. Then, by the same argument as in the proof of Lemma 4, $\sharp\pi_{\mathbb{A}} F_\zeta(\tau) = 2s$ holds for the s-window $\tau := \{0, 1, \cdots, s-1\}$ and $F_\zeta(\tau)$ has at most one constant word. Since $D(\xi, \tau) = 0$ and the window τ satisfies the condition $(*)$ with some η and γ, we have $p_\alpha(\tau) = 2s$ and $\sharp C(F_\alpha(\tau)) \le 1$. This implies that

$$p_\alpha(\tau) - \sharp C(F_\alpha(\tau)) + 2 > 2s,$$

and (2.4) does not holds, which contradicts our assumption. Thus, α is a simple Toeplitz word. ☐

Example 1. Let

$$\begin{aligned} \beta &:= (a?)^\infty \triangleleft (b?)^\infty \triangleleft (a?)^\infty \triangleleft (b?)^\infty \triangleleft \cdots \\ &= abaaabababaaabaa\cdots, \end{aligned}$$

which is a simple Toeplitz word. Let $\xi = (ab?)^\infty$. Then, ξ satisfy (2.7). Therefore, by Theorem 3,

$$\alpha := \xi \triangleleft \beta = abaabbabaabaabaabbabaabbabaabbabaaba\cdots$$

is a pattern Sturmian word. It is not a simple Toeplitz word since $F_\alpha(\{0,1,2\}) = \{aab, aba, baa, abb, bab, bba\}$ and hence,

$$p_\alpha(\{0,1,2\}) - \#C(F_\alpha(\{0,1,2\})) + 2 = 8 > 2 \cdot 3,$$

and (2.4) is not satisfied.

Theorem 4. *Let $\alpha = \xi \lhd \xi \lhd \cdots \in \mathbb{A}^{\mathbb{N}}$ with $\xi \in \mathcal{P}(\mathbb{A}, ?)$. Then, we have*

$$\lim_{k\to\infty} \frac{p^*_\alpha(k)}{k} = \sharp\mathbb{A} + \max_{\substack{L\subset\{0,1,\cdots,r-1\}\\ \sharp L\geq 2}} \frac{E(\xi, L)}{\sharp L - 1},$$

where r is the minimum period of ξ, and $E(\xi, L)$ is defined in (2.3).

Proof. Note that $E(\xi, L) = D(\xi \lhd \xi, rL)$ for any $L \subset \{0, 1, \cdots, r-1\}$.
Let

$$E^* := \max_{\substack{L\subset\{0,1,\cdots,r-1\}\\ \sharp L\geq 2}} \frac{E(\xi, L)}{\sharp L - 1}.$$

Take L_0 with $\sharp L_0 = \ell_0 \geq 2$ attaining the maximum of $E(\xi, L)/(\sharp L - 1)$. Without loss of generality, we may assume that $0 \in L_0$. Let $L'_0 = L_0 \setminus \{0\}$. Define a sequence of windows by $\tau(1) := rL_0$ and

$$\tau(k+1) := r(r\tau(k) \cup L'_0) \quad (k = 1, 2, \cdots).$$

Let us apply (3) of Corollary 1 with $\xi \lhd \xi$ for ξ, α for β, ξ for η, α for η and rL_0 for L. Then, $\tau(k)$ $(k = 1, 2, \cdots)$ satisfy Condition$(*)$. Moreover, $\overline{\tau(k+1)}^0 = \tau(k)$ $(k = 1, 2, \cdots)$. Since $D(\xi \lhd \xi, rL_0) = (\ell_0 - 1)E^*$, we have by (3) of Corollary 1 that

$$\begin{aligned}
p_\alpha(\tau(1)) &= D(\xi \lhd \xi, rL_0) + \sum_{i\in rL_0} p_\alpha(\{i\}) \\
&= (\ell_0 - 1)E^* + \ell\sharp\mathbb{A} \\
p_\alpha(\tau(k+1)) &= D(\xi \lhd \xi, rL_0) + p_\alpha(\tau(k)) + \sum_{i\in rL'_0} p_\alpha(\{i\}) \\
&= (\ell_0 - 1)E^* + p_\alpha(\tau(k)) + (\ell_0 - 1)\sharp\mathbb{A},
\end{aligned}$$

and hence,

$$p_\alpha(\tau(k)) = (\sharp\mathbb{A} + E^*)(K - 1) + \sharp\mathbb{A} \quad (k = 1, 2, \cdots),$$

where $K := k(\ell_0 - 1) + 1$ is the size of the window $\tau(k)$. Thus, we have

$$\liminf_{k\to\infty} p^*_\alpha(k)/k \geq \lim_{k\to\infty} p_\alpha(\tau(k))/(K-1) = \sharp\mathbb{A} + E^*.$$

Now, let us prove that $\limsup_{k\to\infty} p^*_\alpha(k)/k \leq \sharp\mathbb{A} + E^*$. For any $k \geq 2$, let $p^*_\alpha(k) = p_\alpha(\tau')$ for some k-window τ'. By Lemma 1, we may assume that τ' is divisible by r but not by r^2. Hence, we put $\tau' = r\tau$ with τ which is not divisible by

r. We use the decomposition of $r\tau$ given in (2.1) and (2.2) for $\xi \lhd \xi$ instead of ξ. Then by Theorem 1, we have

$$\begin{aligned} p^*_\alpha(k) &= p_\alpha(r\tau) \\ &\leq D(\xi \lhd \xi, rL) + \sum_{i \in L} p_\alpha(r\tau^i) \\ &\leq E(\xi, L) + \sum_{i \in L} p^*_\alpha(k_i) \end{aligned} \tag{2.8}$$

where $k_i = \sharp\tau^i \geq 1$ for any $i \in L$. Since $k = \sum_{i \in L} k_i$ with $\sharp L \geq 2$, any of k_i's for $i \in L$ is less than k. Hence, we can apply the induction on k to prove that

$$p^*_\alpha(k) + E^* \leq (\sharp\mathbb{A} + E^*)k. \tag{2.9}$$

For $k = 1$, (2.9) holds since $p^*_\alpha(k) = \sharp\mathbb{A}$. Let $k \geq 2$. Assume that (2.9) holds for $1, 2, \cdots, k-1$. Then by (2.8) and the assumption of the induction, we have

$$\begin{aligned} p^*_\alpha(k) + E^* &\leq E(\xi, L) + \sum_{i \in L} p^*_\alpha(k_i) + E^* \\ &\leq (\ell - 1)E^* + \sum_{i \in L} p^*_\alpha(k_i) + E^* \\ &= \sum_{i \in L} (p^*_\alpha(k_i) + E^*) \\ &\leq \sum_{i \in L} (\sharp\mathbb{A} + E^*)k_i = (\sharp\mathbb{A} + E^*)k. \end{aligned}$$

Thus, we have (2.9) for $k = 1, 2, \cdots$, and hence, $\limsup_{k\to\infty} p^*_\alpha(k)/k \leq \sharp\mathbb{A} + E^*$, which completes the proof. □

Example 2. Let $\xi = (a^n bba^{n+1}?)^\infty$ and $\alpha = \xi \lhd \xi \lhd \xi \lhd \cdots$ for $n \geq 1$. Then, $E(\xi, L) = 0$ if $\sharp L = 2$, $E(\xi, L) \leq 2$ if $\sharp L \geq 3$ with the equality for $L = \{0, 1, n+2\}$. Hence, we have $\max_{L,\ \sharp L \geq 2} E(\xi, L)/(\sharp L - 1) = 1$. Thus, $\lim_{n\to\infty} p^*_\alpha(k)/k = 3$.

3 Further examples and open problems

By (3) of Corollary 1, we can calculate almost exact values of the maximal pattern complexity $p^*_\alpha(k)$ of the Toeplitz words α just by calculating $D(\xi, L)$.

It is well known that the measure-theoretic dynamical systems arising from our Toeplitz words have discrete spectrum [JK]. On the other hand, it is known [KZ1] that if the maximal pattern complexity of a word increases in less than exponential order, then the measure-theoretic dynamical system arising from it has a discrete spectrum but the converse is not true. Here, we give such examples of Toeplitz words $\alpha \in \mathbb{A}^\mathbb{N}$ with $\sharp\mathbb{A} = 2$ and $p^*_\alpha(k) = 2^k$ ($k = 1, 2, \cdots$). We also give $\alpha \in \mathbb{A}^\mathbb{N}$ with $p^*_\alpha(k)$ increasing in the polynomial order of given degree ≥ 1.

It is known [K] that for any word $\alpha \in \mathbb{A}^{\mathbb{N}}$ with $\sharp\mathbb{A} = 2$, either

$$p^*_\alpha(k) \leq \sum_{i=0}^{n-1} \binom{k}{i} \quad (k = 1, 2, \cdots)$$

for some $n = 1, 2, \cdots$, or

$$p^*_\alpha(k) = 2^k \quad (k = 1, 2, \cdots)$$

holds.

Example 3. Let $\mathbb{A} = \{a, b\}$. For $n = 1, 2, \cdots$, we can take $\eta_n \in \mathbb{A}^n$ with length $n2^n$ such that η_n contains every block in $\mathbb{A}^n$. For example,

$$\eta_1 = ab, \ \eta_2 = aaabbabb, \ \eta_3 = aaaababababaaabbbabbbabbb, \cdots$$

Let $\xi^n = (\eta_n ?)^\infty \in \mathcal{P}(\mathbb{A}, ?)$. Then, we have

$$D(\xi^n, \{0, 1, \cdots, n-1\}) = 2^n - 2n.$$

Let

$$\begin{aligned} &\alpha^k = \xi^k \lhd \xi^{k+1} \lhd \cdots \\ &\tau(k) = \{i\ell_k;\ i = 0, 1, \cdots, k-1\} \quad (k = 1, 2, \cdots), \\ &\text{where } \ell_1 := 1 \text{ and } \ell_k := \prod_{i=1}^{k-1} (i2^i + 1). \end{aligned}$$

Then, by Corollary 1, we have

$$p_\alpha(\tau) = p_{\alpha^k}(\tau(k)/\ell^k) = D(\xi^k, \{0, 1, \cdots, k-1\}) + 2k = 2^k.$$

Thus, $p^*_\alpha(k) = 2^k$ $(k = 1, 2, \cdots)$.

Example 4. (Xue [X]) Let

$$\theta : \begin{array}{ccc} a & \to & aabb \\ b & \to & aaab \end{array}$$

be a substitution on $\mathbb{A} = \{a, b\}$. The fixed point $aabbaabbaaabaaab\cdots$ is a Toeplitz word such that

$$\alpha = (aa?b)^\infty \lhd (bb?a)^\infty (aa?b)^\infty \lhd (bb?a)^\infty \lhd \cdots.$$

Let

$$\xi = (aa?b)^\infty \lhd (bb?a)^\infty = aabbaabba?abaaab.$$

Then, $E(\xi, L) = 0$ if $\sharp L = 2$ or 3, $E(\xi, L) \leq 2$ if $\sharp L \geq 4$ with the equality for $\xi = \{0, 1, 2, 3\}$. Thus, $\max_{L,\ \sharp L \geq 2} E(\xi, L)/(\sharp L - 1) = 2/3$ and $\lim_{n \to \infty} p^*_\alpha(k)/k = 8/3$. This gives an alternative proof of the well known fact (Goodman [G], for example) that the measure-theoretic dynamical system arising from this α has a discrete spectrum since $p^*_\alpha(k)$ increases less than exponentially as the function of k.

Open Problems:

1. What is the maximal pattern complexity of the Toeplitz words $\alpha = \xi \lhd \xi \lhd \xi \lhd \cdots$ with $\xi \in \mathcal{P}(\mathbb{A}, ?)$ having more than one hole in the minimum cycle. Is there a Toeplitz word in this extended class having the maximal pattern complexity increasing exponentially? Cassaigne and Karhumäki [CK] obtained the increasing order k^λ of the block complexity in this class. For example, $\alpha := (ab?a?)^\infty$ has the block complexity increasing in the order $k^{\log 5/(\log 5 - \log 2)}$. The maximal pattern complexity should increase at least as this. We don't know even whether the order is sub-exponential or not.

2. All the pattern Sturmian words known so far except for the simple Toeplitz words fail to satisfy (2.4). Does the property (2.4) characterize the simple Toeplitz words among all the words?

3. Is there a Toeplitz word with one hole such that the maximal pattern complexity increases in a polynomial order with degree > 1?

Acknowledgement: The authors thank Professor Wen Zhi-Ying (Tsinghua University) for inviting 2 of the authors (Kamae, Tan) to Morningside center of Mathematics (CAS) giving them a chance to complete the paper which has been suspended for 3 years. Also thanks go to Professor Rao Hui (Tsinghua University) for helpful discussions.

References

[B] V. Berthé, Sequences of low complexity: automatic and Sturmian sequences, *London Math. Soc. Lecture Note Ser.* 279, Cambridge Univ. Press, 2000, 1-34.

[CK] J. Cassaigne and J. Karhumäki, Toeplitz words, generalized periodicity and periodically iterated morphisms, *European J. Combin.* **18:5** (1997), 497–510.

[G] T.N.T.Goodman, Topological sequence entropy, *Proc. London Math. Soc.* (3) 29 (1974), 331-350.

[JK] K. Jacobs and M. Keane, $0-1$-sequences of Toeplitz type, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **13** (1969), 123-131.

[HY] Wen Huang and Xiangdong Ye, Maximal pattern entropy and null systems, *preprint.*

[K] T. Kamae, Maximal pattern complexity as topological invariants, *preprint.*

[KR] T. Kamae and H. Rao, Pattern complexity over ℓ letters, *to appear in European J. Combinatorics* **27** (2006), 125-137.

[KX] T. Kamae and Y.M. Xue, Two dimensional word with $2k$ maximal pattern complexity, *Osaka J. Math.* **41** (2004), 257-265.

[KZ1] T. Kamae and L. Zamboni, Sequence entropy and the maximal pattern complexity of infinite words, *Ergodic Theory and Dynamical Systems* **22:4** (2002), 1191-1199.

[KZ2] T. Kamae and L. Zamboni, Maximal pattern complexity for discrete systems, *Ergodic Theory and Dynamical Systems* **22:4** (2002), 1201-1214.

[MH] M. Morse and G.A. Hedlund, Symbolic dynamics II: Sturmian sequences, *Amer. J. Math.* **62** (1940), 1-42.

[X] Y.M. Xue, Sequence entropy and maximal pattern complexity, *Doctoral Thesis at Osaka City University*, 2003.