Ergod. Th. & Dynam. Sys., (2026), **46**, 2131–2148 © The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

doi:10.1017/etds.2026.10303

On the structure of sequences with minimal maximal pattern complexity

CASEY SCHLORTT

Department of Mathematics, University of Denver, USA
(e-mail: casey.schlortt@du.edu)

(*Received* 12 *May* 2025 *and accepted in revised form* 17 *March* 2026)

Abstract. In 2002, Kamae and Zamboni [*Ergod. Th. & Dynam. Sys.* **22**(4) (2002), 1191–1199] introduced maximal pattern complexity and determined that any aperiodic sequence must have maximal pattern complexity at least $2k$. In 2006, Kamae and Rao [*European J. Combin.* **27**(1) (2006), 125–137] examined the maximal pattern complexity of sequences over larger alphabets and showed that sequences which have maximal pattern complexity less than ℓk, for ℓ the size of the alphabet, must have some periodic structure. In this paper, we investigate the structure of sequences of low maximal pattern complexity over ℓ letters, where $\liminf_{k\to\infty} p^*_\alpha(k) - 3k = -\infty$. In addition, we show that the minimal maximal pattern complexity of an aperiodic sequence which uses all ℓ letters is $p^*_\alpha(k) = 2k + \ell - 2$ and give an exact structure for aperiodic sequences with this maximal pattern complexity.

Key words: symbolic dynamics, maximal pattern complexity, pattern Sturmian sequences, strong pattern Sturmian sequences
2020 Mathematics Subject Classification: 37B10 (Primary)

1. *Introduction*

There are various methods to study the complexity of symbolically defined dynamical systems. Block (word) complexity, which counts the number of blocks of size k which appear in a subshift, was introduced in 1938 by Morse and Hedlund as a way to determine the 'size' of different subshifts. Their first paper, [**MH38**], contains the proof of the Morse–Hedlund theorem, which states that any aperiodic sequence must have block complexity at least $k + 1$ for all $k \in \mathbb{N}$. Their second paper, [**MH40**], introduces *Sturmian sequences*, defined to be infinite sequences over two letters that have block complexity $k + 1$ for all $k \in \mathbb{N}$ and that can be represented as a coding of an irrational circle rotation.

For larger alphabets, a similar proof shows that the block complexity of an aperiodic sequence over $\ell \geq 2$ letters must be at least $k + \ell - 1$. A *Sturmian sequence over* $\ell \geq 2$ *letters* is defined to be an aperiodic sequence with block complexity $k + \ell - 1$ for all $k \in \mathbb{N}$ [**Fog02**].

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

The maximal pattern complexity of an infinite sequence α, denoted $p^*_\alpha(k)$, was introduced by Kamae and Zamboni in 2002 in [KZ02b] as a way to link word complexity and sequence entropy. In their first paper, they showed that the maximal pattern complexity of any aperiodic sequence must be at least $2k$ for all $k \in \mathbb{N}$, which is a direct analog to the Morse–Hedlund theorem.

In [KZ02b], they also introduced *pattern Sturmian sequences* as an analog to Sturmian sequences and defined pattern Sturmian sequences over two letters to be aperiodic sequences with the lowest possible maximal pattern complexity, $2k$. In [KZ02a], they furthered their study of pattern Sturmian sequences, showing that sequences generated by a coding of an irrational circle rotation by any two interval partitions of a circle and simple Toeplitz sequences are examples of pattern Sturmian sequences over two letters.

In 2006, Kamae and Rao extended the study of maximal pattern complexity for sequences over finite alphabets of size greater than two in [KR06]. They determined that a sequence over $\ell \geq 2$ letters which has maximal pattern complexity less than ℓk for some $k \in \mathbb{N}$ is *periodic by projection*, meaning there is a non-constant, coordinatewise, factor map f that maps the sequence to a periodic sequence. They extended the definition of pattern Sturmian sequences over two letters to *pattern Sturmian sequences over* $\ell \geq 2$ *letters*, defining them to be sequences of maximal pattern complexity ℓk that are not periodic by projection.

As stated in their paper, their results hold for sequences in which every letter appears infinitely often. We will make this same assumption for all sequences in this paper, since the ability for a letter to show up a finite number of times can be eliminated with a finite shift.

Although [KR06] isolates the minimal maximal pattern complexity for aperiodic sequences which are not periodic by projection, there are aperiodic sequences with lower maximal pattern complexity. The goal of this work is to isolate the lowest maximal pattern complexity for any aperiodic sequence and determine the structure of such sequences. We define a *strong pattern Sturmian sequence* over $\ell \geq 2$ letters to be an aperiodic sequence with maximal pattern complexity $2k + \ell - 2$. As seen in Theorem 2.11, this is the lowest possible maximal pattern complexity an aperiodic sequence can have (given all letters of the alphabet appear in the sequence) and this notion aligns closely with the definition of a Sturmian sequence over ℓ letters. Furthermore, since the notions of pattern Sturmian sequences and strong pattern Sturmian sequences coincide when the alphabet is two letters, we will restrict our study to sequences for which the size of the alphabet is greater than two.

In this paper, we introduce three main results. The first result, Theorem 1.1, gives a general structure for aperiodic sequences with low maximal pattern complexity. Example 6.3 demonstrates that the bound of $4k$ in Theorem 1.1 cannot be substantially improved.

THEOREM 1.1. *Let α be an aperiodic sequence such that $\#\mathcal{A}_\alpha = \ell \geq 3$. If*

$$\liminf_{k\to\infty} p^*_\alpha(k) - 3k = -\infty, \tag{1.1}$$

then there exists some decomposition of α into $m > 1$ residues

$$\{\alpha^{(i)} = \alpha_i \alpha_{i+m} \alpha_{i+2m} \ldots : 0 \leq i \leq m-1\}$$

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

such that exactly one of these residues is aperiodic over eventually only two letters, a and b, and any residue containing finitely many a and b symbols is eventually constant.

If α is uniformly recurrent, then the bound in (1.1) can be improved to $4k$.

The second result, Theorem 1.2, states that any strong pattern Sturmian sequence must consist of a pattern Sturmian sequence over two letters threaded together with constant sequences. Example 6.1 demonstrates the existence of sequences which are strong pattern Sturmian.

THEOREM 1.2. *Let α be an aperiodic sequence such that $\#\mathcal{A}_\alpha = \ell \geq 3$. If $p^*_\alpha(k) = 2k + \ell - 2$ for all $k \in \mathbb{N}$, then the aperiodic residue from Theorem 1.1 must be a pattern Sturmian over two letters and all other residues must be constant.*

The third result, Theorem 1.3, is a partial converse of Theorem 1.2, which states that a sequence that can be decomposed into a pattern Sturmian sequence threaded together with constant sequences must have maximal pattern complexity $2k + C$ for some constant $C \in \mathbb{N}$. Example 6.2 demonstrates that a literal converse to Theorem 1.2 is untrue. When $C = \ell - 2$, the sequence is strong pattern Sturmian.

THEOREM 1.3. *Let α be an aperiodic sequence such that $\#\mathcal{A}_\alpha = \ell \geq 3$. If α can be decomposed into residues such that one residue is a pattern Sturmian over two letters and all other residues are constant, then for sufficiently large k, $p^*_\alpha(k) = 2k + C$ for some $C \in \mathbb{N}$ with $C \geq \ell - 2$.*

It is important to note that Theorem 1.1 allows for certain residues to be eventually periodic over two letters, but any non-constant periodic residue must be periodic over the alphabet of the aperiodic residue. However, Theorem 1.2 requires that any residue be aperiodic or constant. Strong pattern Sturmian sequences also cannot have eventually constant residues and the aperiodic residue can only witness two letters.

It is also important to recognize that in the case where $\ell = 3$, these results give a general structure for all aperiodic sequences α that have maximal pattern complexity less than $3k$. These sequences require a single aperiodic residue that is eventually over two letters, a and b, threaded together with eventually constant sequences of the third letter c.

2. *Definitions and preliminaries*

This section is dedicated to introducing definitions and results needed in the following sections.

Definition 2.1. Let $\mathcal{A}$ be a finite set and $\alpha = \alpha_0\alpha_1\alpha_2 \ldots \in \mathcal{A}^{\mathbb{N}_0}$. Then, α is a *(one-sided) infinite sequence over finite alphabet* $\mathcal{A}$.

As mentioned in the introduction, we only consider sequences in which a letter appearing once in the sequence means the letter appears infinitely in the sequence.

Throughout this paper, $\mathcal{A}^{\mathbb{N}_0}$ is endowed with the product topology.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

Definition 2.2. Let $\sigma : \mathcal{A}^{\mathbb{N}_0} \to \mathcal{A}^{\mathbb{N}_0}$ be the *shift map* defined by $\sigma(\alpha_0\alpha_1\alpha_2\ldots) = \alpha_1\alpha_2\ldots$. We say that a *finite shift of* α is the sequence $\sigma^t(\alpha) = \alpha_t\alpha_{t+1}\alpha_{t+2}\ldots$ for any integer $t \in \mathbb{N}_0$. We say that the *orbit* of α, denoted $\mathcal{O}(\alpha)$, is the set $\mathcal{O}(\alpha) := \{\sigma^t(\alpha) : t \in \mathbb{N}_0\}$. The *orbit closure* of α is the orbit of α together with its limit points and is denoted $\overline{\mathcal{O}(\alpha)}$.

While α uses letters from $\mathcal{A}$, it may not use all letters from $\mathcal{A}$. Thus, to denote the letters that appear in an infinite sequence α, we use

$$\mathcal{A}_\alpha := \{a \in \mathcal{A} : \text{ there exists } i \in \mathbb{N}_0 \text{ such that } \alpha_i = a\}.$$

Definition 2.3. A sequence α is *periodic* if there exists some $t \in \mathbb{N}$ such that $\alpha = \sigma^t(\alpha)$. A sequence α is *eventually periodic* if there exists some $t \in \mathbb{N}$ such that $\sigma^t(\alpha)$ is periodic. A sequence α is *aperiodic* if α is not periodic or eventually periodic.

Definition 2.4. A sequence α is *recurrent* if for all $L \in \mathbb{N}$, there exists some $M \in \mathbb{N}$ such that $\alpha_0 \ldots \alpha_{L-1} = \alpha_M \ldots \alpha_{M+L-1}$. If for all $L \in \mathbb{N}$, the set $R_L = \{M : \alpha_0 \ldots \alpha_{L-1} = \alpha_M \ldots \alpha_{M+L-1}\}$ is syndetic (there are bounded gaps between subsequent elements), then α is *uniformly recurrent*.

Definition 2.5. A *k-sized window* τ is an increasing set of k indices beginning with zero:

$$\tau := \{0 = \tau(0) < \tau(1) < \cdots < \tau(k-1)\} \subset \mathbb{N}_0.$$

We say a $(k+1)$-sized window τ' is an *immediate extension* of a k-sized window τ if $\tau' = \{\tau(0), \tau(1), \ldots, \tau(k-1), h\}$ for some $h > \tau(k-1)$.

Given a k-sized window τ, we use the notation $m\tau$ to denote the k-sized window $m\tau := \{m\tau(0), m\tau(1), \ldots, m\tau(k-1)\}$.

Definition 2.6. A *τ-word over* α is a finite word of length k denoted $\alpha[n+\tau]$ and defined as

$$\alpha[n+\tau] := \alpha_{\tau(0)+n}\alpha_{\tau(1)+n}\ldots\alpha_{\tau(k-1)+n} \in \mathcal{A}^k$$

for some $n \in \mathbb{N}_0$.

The *set of all τ-words over* α is the set denoted by $F_\alpha(\tau)$ and defined by

$$F_\alpha(\tau) := \{\alpha[n+\tau] : n \in \mathbb{N}_0\}.$$

Since $F_\alpha(\tau)$ is finite, there exists some $L \in \mathbb{N}$ such that $F_\alpha(\tau) = \{\alpha[n+\tau] : n \in \mathbb{N} \cap [0, L)\}$.

Definition 2.7. The *maximal pattern complexity* of an infinite sequence α is the function $p^*_\alpha : \mathbb{N} \to \mathbb{N}$ defined by

$$p^*_\alpha(k) := \sup_\tau \#F_\alpha(\tau),$$

where the supremum is taken over all k-sized windows τ.

We note that for any $k \in \mathbb{N}$, there exists some k-sized window τ such that $p^*_\alpha(k) = \#F_\alpha(\tau)$ and we call any such window τ a *maximal k-sized window* τ.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

In [KZ02b], Kamae and Zamboni proved the following two theorems for maximal pattern complexity, the first that acts as an analog to the Morse–Hedlund theorem for block complexity and the second that relates the number of τ-words to the number of τ'-words in an aperiodic sequence α for some immediate extension τ' of τ.

THEOREM 2.8. [KZ02b] *An infinite sequence $\alpha \in \mathcal{A}^{\mathbb{N}_0}$ for a finite alphabet $\mathcal{A}$ is eventually periodic if and only if $p^*_\alpha(k) \leq 2k - 1$ for some $k \in \mathbb{N}$.*

THEOREM 2.9. [KZ02b] *For an arbitrary word α over a finite set $\mathcal{A}$, if there exists $k \in \mathbb{N}$ and a window τ of size k such that $\#F_\alpha(\tau') \leq \#F_\alpha(\tau) + 1$ holds for any immediate extension τ' of τ, then α is eventually periodic.*

The following fact is an extension of the Morse–Hedlund theorem for any window τ, which follows from the same proof.

Fact 2.10. For any aperiodic sequence α, any $k \in \mathbb{N}$, and any k-sized window τ, $\#F_\alpha(\tau) \geq k + 1$.

Using Theorem 2.9, it is easy to extend Theorem 2.8 to the case where $\mathcal{A}_\alpha = \mathcal{A}$ as shown in the following theorem.

THEOREM 2.11. *If an infinite sequence α is aperiodic and $\#\mathcal{A}_\alpha = \ell \geq 2$, then $p^*_\alpha(k) \geq 2k + \ell - 2$ for all $k \in \mathbb{N}$.*

Proof. Suppose α is aperiodic and $\#\mathcal{A}_\alpha = \ell \geq 2$. Clearly, since $\#\mathcal{A}_\alpha = \ell$, $p^*_\alpha(1) = \ell$.

Suppose for some $k > 1$, we have $p^*_\alpha(k) \geq 2k + \ell - 2$.

Let τ be a maximal k-sized window for α. Then, $\#F_\alpha(\tau) \geq 2k + \ell - 2$. By Theorem 2.9, since α is aperiodic, there exists some immediate extension τ' of τ such that $\#F_\alpha(\tau') \geq \#F_\alpha(\tau) + 2$. Thus,

$$p^*_\alpha(k+1) \geq \#F_\alpha(\tau') \geq \#F_\alpha(\tau) + 2 \geq 2(k+1) + \ell - 2.$$

Hence, for all $k \in \mathbb{N}$, $p^*_\alpha(k) \geq 2k + \ell - 2$. □

In [KZ02b], Kamae and Zamboni defined the maximal pattern complexity analog for a Sturmian sequence over a two-letter alphabet.

Definition 2.12. A *pattern Sturmian sequence over a two-letter alphabet* is an aperiodic sequence α for which $p^*_\alpha(k) = 2k$ for all $k \in \mathbb{N}$.

Kamae and Rao introduced the singular decomposition of an infinite sequence in [KR06]. In what follows is a quick outline of this process, but a more in depth exploration can be found in [KR06].

In Definition 2.15, the singular decomposition of a recurrent sequence is defined and in Definition 2.17, the singular decomposition of a non-recurrent sequence is defined. A singular decomposition requires a recurrent sequence, so for a non-recurrent sequence α, we need to find some recurrent sequence β in the orbit closure of α that satisfies some technical conditions outlined in Definition 2.16, then decompose α with respect to the

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

decomposition of β. The following definitions outline the singular decomposition of any sequence α.

Definition 2.13. A sequence α is *periodic by projection* if there exists some proper, non-empty subset of the alphabet, $S \subset \mathcal{A}$, such that the sequence $1_S(\alpha_0)1_S(\alpha_1)1_S(\alpha_2)\ldots \in \{0,1\}^{\mathbb{N}_0}$ is eventually periodic.

A letter $b \in \mathcal{A}_\alpha$ is *singular* if the sequence $1_{\{b\}}(\alpha_0)1_{\{b\}}(\alpha_1)1_{\{b\}}(\alpha_2)\ldots \in \{0,1\}^{\mathbb{N}_0}$ is periodic.

Definition 2.14. Let β be some recurrent sequence. We say $m \in \mathbb{N}$ is the *decomposition cycle of* β if m is the smallest integer which is a period of $1_{\{b\}}(\alpha_0)1_{\{b\}}(\alpha_1)1_{\{b\}}(\alpha_2)\ldots$ for each singular letter b of β. If β has no singular letters, then we set $m = 1$.

Definition 2.15. The *singular decomposition of a recurrent sequence* β is the set of sequences

$$\{\beta^{(i)} = \beta_i\beta_{i+m}\beta_{i+2m}\ldots : 0 \le i < m\},$$

where m is the decomposition cycle of β.

Definition 2.16. Choose some $r \in \{0,1,\ldots,m-1\}$ such that $\lim_{n\to\infty}\sigma^{mk_n+r}(\alpha)$ converges for some sequence $(k_n)_{n\in\mathbb{N}}$. Let $\beta = \sigma^{m-r}(\lim_{n\to\infty}\sigma^{mk_n+r}(\alpha))$. Here, β is an *auxiliary word of* α if the following conditions are true:

(1) β is recurrent;
(2) $\beta \in \overline{\mathcal{O}(\alpha)}$;
(3) $\beta^{(i)} \in \overline{\mathcal{O}(\alpha^{(i)})}$ for all $i \in \mathbb{N} \cap [0, m-1]$.

In [**KR06**], Kamae and Rao proved that any aperiodic sequence has at least one auxiliary word.

Note that if β is an auxiliary word of α, then $p^*_\alpha(k) \ge p^*_\beta(k)$ for all $k \in \mathbb{N}$ since for any k-sized window τ, $F_\beta(\tau) \subseteq F_\alpha(\tau)$.

Definition 2.17. The *singular decomposition of a non-recurrent word* α *with respect to an auxiliary word* β is the set of sequences

$$\{\alpha^{(i)} = \alpha_i\alpha_{i+m}\alpha_{i+2m}\ldots : 0 \le i < m\}$$

for m the decomposition cycle of the auxiliary word β.

Definition 2.18. Define an undirected graph Γ on the residues of the singular decomposition of α by $V(\Gamma) = \{0,1,2,\ldots,m-1\}$, the residues of α, and $E(\Gamma) = \{(i,j) : \mathcal{A}_{\alpha^{(i)}} \cap \mathcal{A}_{\alpha^{(i)}} \ne \emptyset\}$.

Definition 2.19. For any infinite sequence α, k-sized window τ, and residue i, $0 \le i \le m-1$ for m the decomposition cycle of α, we define the *set of* τ*-words beginning on the* i*th residue* as the set

$$C_i(\tau,\alpha) := \{\alpha[n+\tau] : n \equiv i \mod m\}.$$

We note that $\bigcup_{i=0}^{m-1} C_i(\tau,\alpha) = F_\alpha(\tau)$.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

In [KR06], Kamae and Rao also proved the following theorems about the maximal pattern complexity of sequences and found a threshold under which a non-trivial singular decomposition is forced.

THEOREM 2.20. [KR06] *Let β be a recurrent sequence over ℓ letters, $\ell \geq 2$. If $p^*_\beta(k) < \ell k$ holds for some $k \in \mathbb{N}$, then β contains at least one singular letter.*

THEOREM 2.21. [KR06] *If $\ell \geq 2$ and $p^*_\alpha(k) < \ell k$ holds for some $k \in \mathbb{N}$, then Γ is disconnected.*

Definition 2.22. [KR06] A sequence α over $\ell \geq 2$ letters is called a *pattern Sturmian sequence* if $p^*_\alpha(k) = \ell k$ and α is not periodic by projection.

We define the following slightly more restrictive condition.

Definition 2.23. A sequence α over $\ell \geq 2$ letters is called a *strong pattern Sturmian sequence* if $p^*_\alpha(k) = 2k + \ell - 2$.

Clearly, if $\ell = 2$, the notions of pattern Sturmian and strong pattern Sturmian are the same.

In general, two sequences may have very different maximal windows. However, when both sequences are uniformally recurrent, the following theorem shows we can find a window which gives at least $2k$ words for each sequence. This is a shared maximal window when both sequences are pattern Sturmian. We prove this with an argument from the proof of Theorem 2.9, which can be found in [KZ02b].

THEOREM 2.24. *For any two aperiodic sequences α and β that are uniformly recurrent and for each $k \in \mathbb{N}$, there exists some k-sized window τ such that $\#F_\alpha(\tau) \geq 2k$ and $\#F_\beta(\tau) \geq 2k$.*

Proof. Consider the product space $\overline{\mathcal{O}(\alpha)} \times \overline{\mathcal{O}(\beta)}$ with the map $\sigma \times \sigma$. $\overline{\mathcal{O}(\alpha)} \times \overline{\mathcal{O}(\beta)}$ has a recurrent point (by a simple application of the Poincaré recurrence theorem), call it (γ, η).

Since α and β are uniformly recurrent, γ and η are both aperiodic.

Clearly, γ and η share a window of size one, and since they are both aperiodic, this window gives at least two words of length one for each sequence.

Now, suppose for some $k > 1$, there exists a $(k-1)$-sized window τ such that $F_\gamma(\tau) \geq 2(k-1)$ and $F_\eta(\tau) \geq 2(k-1)$.

Let $L \in \mathbb{N}$ such that $\gamma_0\gamma_1 \ldots \gamma_{L-1}$ and $\eta_0\eta_1 \ldots \eta_{L-1}$ witness all τ words over γ and η, respectively. Since (γ, η) is recurrent, there exists $n \in \mathbb{N}$ such that for all $0 \leq i < L$, $(\gamma, \eta)_i = (\gamma, \eta)_{n+i}$. Therefore, for all $0 \leq i < L$, $\gamma_i = \gamma_{i+n}$ and $\eta_i = \eta_{i+n}$.

Define τ' to be an immediate extension of τ by $\tau'(i) = \tau(i)$ for all $0 \leq i \leq k-2$ and $\tau'(k-1) = n$.

In the proof of Theorem 2.9 from [KZ02b], the following fact was proved: if δ is aperiodic and there is some $\tau' = \tau \cup \{n\}$ for some n such that $awa \in \#F_\delta(\tau')$ for all $aw \in \#F_\delta(\tau)$, a a single letter, and w a $k-2$-sized $\tau \setminus \{0\}$ word, then $\#F_\delta(\tau') \geq \#F_\delta(\tau) + 2$.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

Thus, for $\delta \in \{\gamma, \eta\}$, $\#F_\delta(\tau') \geq \#F_\delta(\tau) + 2$, and by the inductive hypothesis, $\#F_\alpha(\tau') \geq 2k$ and $\#F_\beta(\tau') \geq 2k$. □

3. *Recurrent case*

The purpose of this section is to prove Theorem 1.1 in the case where α is recurrent.

For the entirety of this section, we assume α is aperiodic, recurrent, $\#\mathcal{A}_\alpha = \ell \geq 3$, and $\liminf_{k\to\infty} p^*_\alpha(k) - 3k = -\infty$ (or in the case where α is uniformly recurrent, that $\liminf_{k\to\infty} p^*_\alpha(k) - 4k = -\infty$).

By Theorem 2.20, since $p^*_\alpha(k) < 3k \leq \ell k$ for some $k \in \mathbb{N}$, α has a singular letter. Let m be the decomposition cycle of α, $m \geq 2$ since the decomposition is non-trivial by the existence of a singular letter.

THEOREM 3.1. *If α is aperiodic, recurrent, and $\liminf_{k\to\infty} p^*_\alpha(k) - 3k = -\infty$, then exactly one residue in the singular decomposition of α is aperiodic. In the case where α is uniformly recurrent, the implication holds for $4k$ instead of $3k$.*

Proof. Suppose α is aperiodic, recurrent, and $\liminf_{k\to\infty} p^*_\alpha(k) - 3k = -\infty$. Let m be the decomposition cycle of α, $m \geq 2$.

Since α is aperiodic, there exists some p, $0 \leq p \leq m-1$, such that $\alpha^{(p)}$ is aperiodic. Let $k > m$ and let τ be a maximal $(k-(m-1))$-sized window over $\alpha^{(p)}$. Then, $\#F_{\alpha^{(p)}}(\tau) \geq 2k - 2(m-1)$ by Theorem 2.8.

Define a k-sized window τ' over α by $\tau' = m\tau \cup \{h_i : 1 \leq i \leq m-1\}$ for $h_1 > m\tau(k-(m-1)-1)$, $h_{i+1} > h_i$ for all $1 \leq i \leq m-2$, and $h_i \equiv i \mod m$ for all $1 \leq i \leq m-1$.

Define $x \in (\mathcal{A} \cup \{?\})^{\mathbb{N}}$ by $x_i = \alpha_i$ if α_i is a singular letter and $x_i = ?$ if not. Then, x has least period m since a smaller least period for x would result in the singular letters of α sharing a period smaller than m.

Therefore, $C_i(\tau', \alpha) \cap C_j(\tau', \alpha) = \emptyset$ for all $0 \leq i < j \leq m-1$ since the position of the singular letters uniquely determines to which set $C_i(\tau', \alpha)$ a word belongs. Furthermore, $\#C_p(\tau', \alpha) \geq 2k - 2(m-1)$ since every word in $F_{\alpha^{(p)}}(\tau)$ must appear as a prefix for some word in $C_p(\tau', \alpha)$.

Suppose toward a contradiction that there exists an i, $0 \leq i \leq m-1$ and $i \neq p$, such that $\alpha^{(i)}$ is also aperiodic. Consider τ over $\alpha^{(i)}$. Since $\alpha^{(i)}$ is aperiodic, then by Fact 2.10, $\#F_{\alpha^{(i)}}(\tau) \geq k - (m-1) + 1$. Thus, $\#C_i(\tau', \alpha) \geq k - (m-1) + 1$, and since $C_p(\tau', \alpha)$ and $C_i(\tau', \alpha)$ are disjoint, $\#F_\alpha(\tau') \geq 3k - 3(m-1) + 1$. Therefore,

$$\liminf_{k\to\infty} p^*_\alpha(k) - 3k \geq \liminf_{k\to\infty} 3k - 3(m-1) + 1 - 3k \geq -3(m-1) + 1.$$

Since m is fixed, this is a contradiction.

In the case where α is uniformly recurrent, any aperiodic residue must also be uniformly recurrent. This is because the appearance of singular letters in any block of α forces the points where the block begins again to fall on exactly the same residue as the original block. Thus, blocks in each residue must also appear syndetically, as they directly correspond to the uniform recurrence of the original sequence. Therefore, if $\alpha^{(p)}$ and $\alpha^{(i)}$ are both aperiodic, then they are both uniformly recurrent, and by Theorem 2.24, there exists

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

a $(k-(m-1))$-sized window τ such that $\#F_{\alpha^{(p)}}(\tau) \geq 2k-2(m-1)$ and $\#F_{\alpha^{(i)}}(\tau) \geq 2k-2(m-1)$. Define a k-sized window τ' over α by $\tau' = m\tau \cup \{h_i : 1 \leq i \leq m-1\}$ for $h_1 > m\tau(k-(m-1)-1)$, $h_{i+1} > h_i$ for all $1 \leq i \leq m-2$, and $h_i \equiv i \mod m$ for all $1 \leq i \leq m-1$. Then, $C_p(\tau', \alpha) \cap C_i(\tau', \alpha) = \emptyset$, $\#C_p(\tau', \alpha) \geq 2k-2(m-1)$, and $\#C_i(\tau', \alpha) \geq 2k-2(m-1)$. Therefore,

$$\liminf_{k\to\infty} p^*_\alpha(k) - 4k \geq \liminf_{k\to\infty} 4k - 4(m-1) - 4k \geq -4(m-1),$$

which is a contradiction since m is fixed.

Thus, exactly one residue of the singular decomposition of α is aperiodic. □

PROPOSITION 3.2. *If α is aperiodic, recurrent, and $\liminf_{k\to\infty} p^*_\alpha(k) - 3k = -\infty$, then the aperiodic residue from Theorem 3.1 must be over two letters and for all $c \in \mathcal{A}$ if c is not a letter in the aperiodic residue, c is singular for α.*

Proof. Let $\alpha^{(p)}$ be the aperiodic residue from Theorem 3.1, $0 \leq p \leq m-1$. Suppose toward a contradiction that $\#\mathcal{A}_{\alpha^{(p)}} \geq 3$. Since for any $k \in \mathbb{N}$ and k-sized window τ over $\alpha^{(p)}$, $F_{\alpha^{(p)}}(\tau) \subset F_\alpha(m\tau)$, then $p^*_{\alpha^{(p)}}(k) \leq p^*_\alpha(k) < 3k$. Thus, by Theorem 2.20, $\alpha^{(p)}$ must have a singular letter, call it c. Since all other residues are periodic and c appears in $\alpha^{(p)}$ periodically, c must appear in α periodically, and so c is singular for α and thus cannot appear in $\alpha^{(p)}$, which is a contradiction. Thus, $\#\mathcal{A}_{\alpha^{(p)}} = 2$.

Suppose $c \in \mathcal{A}$, $c \notin \mathcal{A}_{\alpha^{(p)}}$. Then, c only appears in residues which are periodic. Thus, c appears periodically in α, so c is singular for α. □

Clearly, Theorem 3.1 and Proposition 3.2 together prove Theorem 1.1 in the case when α is recurrent.

4. *Non-recurrent case*

This section is dedicated to proving Theorem 1.1 in the non-recurrent case. For the entirety of this section, we assume that α is a non-recurrent, aperiodic sequence, $\#\mathcal{A}_\alpha = \ell \geq 3$, and $\liminf_{k\to\infty} p^*_\alpha(k) - 3k = -\infty$.

Let β be an auxiliary word for α. By Theorem 2.21, since $p^*_\alpha(k) < 3k \leq \ell k$ for some $k \in \mathbb{N}$, Γ is disconnected. Let m be the decomposition cycle of the singular decomposition of α with respect to β, $m \geq 2$ since the decomposition is non-trivial.

For each eventually periodic $\alpha^{(i)}$, there exists an $n_i \in \mathbb{N}_0$ such that $\sigma^{n_i}(\alpha^{(i)})$ is periodic. For each aperiodic $\alpha^{(i)}$, there exists an $n_i \in \mathbb{N}_0$ such that each letter in $\mathcal{A}_{\sigma^{n_i}(\alpha^{(i)})}$ appears in $\sigma^{n_i}(\alpha^{(i)})$ infinitely often. Let $t = \max\{n_i : 0 \leq i \leq m-1\}$. Then, for the sequence $\alpha' = \sigma^{mt}(\alpha)$, all residues of α' are aperiodic or periodic, every letter appearing in a residue of α' appears in that residue infinitely often, $\alpha = w\alpha'$ for some finite word w, and $p^*_\alpha(k) = p^*_{\alpha'}(k) + C'$ for some constant $C' \in \mathbb{N}_0$ and k sufficiently large (since for any window, a finite prefix can only contribute a finite number of words). We further note that β is an auxiliary word for α' and the singular decomposition of α' is $\{\alpha'^{(i)} = \sigma^t(\alpha^{(i)}) : 0 \leq i \leq m-1\}$. The graph Γ' of the residues of α' is such that $V(\Gamma) = V(\Gamma')$ and $E(\Gamma') \subset E(\Gamma)$, and so Γ' is also disconnected.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

LEMMA 4.1. *Exactly one connected component of* Γ' *contains aperiodic residues and all other connected components of* Γ' *contain only singular residues.*

Proof. Since α' is aperiodic, at least one connected component of Γ' contains residues which are aperiodic.

Suppose toward a contradiction that there are two disjoint connected components of Γ' which contain aperiodic residues. Then, there are two residues, $i \neq j$ for which $\alpha'^{(i)}$ and $\alpha'^{(j)}$ are both aperiodic on disjoint alphabets. Let $k \in \mathbb{N}$ and let τ be a maximal k-sized window over $\alpha'^{(i)}$. Then, by Theorem 2.8, $\#F_{\alpha'^{(i)}}(\tau) \geq 2k$. Furthermore, $F_{\alpha'^{(i)}}(\tau) \cap F_{\alpha'^{(j)}}(\tau) = \emptyset$ since the alphabets of $\alpha'^{(i)}$ and $\alpha'^{(j)}$ are disjoint, and by Fact 2.10, $\#F_{\alpha'^{(j)}}(\tau) \geq k+1$ since $\alpha'^{(j)}$ is aperiodic. Then, $\#F_{\alpha'}(m\tau) \geq 2k+k+1$, which is a contradiction.

Let $A \subset \{0, 1, \ldots, m-1\}$ be the set of residues in the aperiodic connected component and $S = \{0, 1, \ldots, m-1\} \setminus A$ be all other residues.

Let $i \in S$. Then, $\alpha'^{(i)}$ is periodic and so $\beta^{(i)}$ is also periodic. Since the letters appearing in $\beta^{(i)}$ only appear in residues from S, those letters must appear periodically in β and thus are singular for β. Therefore, since the letters are singular for β and periodic in α, they must be singular for α', and for all $i \in S$, $\alpha'^{(i)}$ is constant. □

For the remainder of this section, let A be the set of residues in the aperiodic connected component and S be the collection of all other residues.

LEMMA 4.2. *If more than two letters appear in a single aperiodic residue,* $\alpha'^{(i)}$*, then one of the letters must appear periodically within* $\alpha'^{(i)}$.

Proof. Suppose $\alpha'^{(i)}$ is an aperiodic residue for α' and $\#\mathcal{A}_{\alpha^{(i)}} \geq 3$. Then, $p^*_{\alpha'^{(i)}}(k) \leq p^*_{\alpha'}(k) < 3k$ for sufficiently large k.

If $\alpha'^{(i)}$ is recurrent, then by Theorem 2.20, it must have some singular letter and, by definition, this letter is periodic in $\alpha'^{(i)}$.

If $\alpha'^{(i)}$ is not recurrent, then by Theorem 2.21, the graph Γ'_i of $\alpha'^{(i)}$ is disconnected, and by Lemma 4.1, only one connected component has aperiodic pieces and all other pieces are singular for $\alpha'^{(i)}$. Thus, there is at least one letter that appears periodically in $\alpha'^{(i)}$. □

LEMMA 4.3. *If* $c \in \mathcal{A}_A := \bigcup_{i \in A} \mathcal{A}_{\alpha'^{(i)}}$*, then there exists some residue* j*,* $0 \leq j \leq m-1$ *such that* c *appears aperiodically in* $\alpha'^{(j)}$.

Proof. Let $c \in \mathcal{A}_A$. If c appears periodically in each $\alpha'^{(j)}$, $0 \leq j \leq m-1$, then c would appear periodically in each $\beta^{(j)}$. Thus, c would be singular for β and since c is periodic in every residue of α', c would be singular for α', which is a contradiction. □

THEOREM 4.4. *The alphabet of the aperiodic residues,* $\mathcal{A}_A = \bigcup_{i \in A} \mathcal{A}_{\alpha'^{(i)}}$*, has exactly two letters.*

Proof. Clearly, $\#\mathcal{A}_A \geq 2$ or else no residue from A would be aperiodic. Thus, suppose toward a contradiction that $\#\mathcal{A}_A > 2$.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

Since β is contained in the orbit closure of α', $p^*_\beta(k) \leq p^*_{\alpha'}(k)$. Thus, since β is recurrent, β either has the form outlined in the conclusion of Theorem 1.1 by §3, or β is periodic.

Let a, b, and c be distinct letters in $\mathcal{A}_A$ and let $k > m$.

Case I. Suppose β has the form outlined in the conclusion of Theorem 1.1.

Suppose $\beta^{(p)}, 0 \leq p \leq m-1$, is the only aperiodic residue and is over the letters $\{a, b\}$, and all other residues of β are either non-constant periodic over $\{a, b\}$ or constant.

Let τ be a maximal k-sized window over $\alpha'^{(p)}$. Since β is in the orbit closure of α', $F_{\beta^{(p)}}(\tau) \subset F_{\alpha'^{(p)}}(\tau)$ and since $\beta^{(p)}$ is aperiodic, $F_{\alpha'^{(p)}}(\tau)$ contains at least $2k$ τ-words which only have a and b as letters.

By Lemma 4.3, since $c \in \mathcal{A}_A$, there exists some residue i, $0 \leq i \leq m-1$, such that c appears aperiodically (and infinitely) in $\alpha'^{(i)}$. Since $c \neq a, b$, then either $\beta^{(i)} = c^\infty$ or c does not appear in $\beta^{(i)}$ because any non-constant residue of β can only be over the letters a and b.

If $\beta^{(i)} = c^\infty$, then by the aperiodicity of $\alpha'^{(i)}$, there are $k-2$ τ-words over $\alpha'^{(i)}$ of the form $c^j d_j u_j$, $0 < j < k$, $d_j \neq c$, and $u_j \in \mathcal{A}^{k-j-1}$. Thus, $p^*_\alpha(k) \geq \#F_{\alpha'}(m\tau) \geq 3k-2$, which is a contradiction.

If c does not appear in $\beta^{(i)}$, then there are at least $k-2$ τ-words over $\alpha'^{(i)}$ with the form $u_j c u'_j$ for $0 < j < k$, and some $u_j \in (\mathcal{A} \setminus \{c\})^j$, $u'_j \in \mathcal{A}^{k-j-1}$. Thus, $p^*_\alpha(k) \geq \#F_{\alpha'}(m\tau) \geq 3k-2$ for all $k > m$, which is a contradiction.

Case II. Suppose β is periodic.

Since every letter of $\mathcal{A}_A$ must appear aperiodically in some residue of α' and $\#\mathcal{A}_A \geq 3$, by Lemma 4.2, there must be at least two residues of α' which are aperiodic, $\alpha'^{(p)}$ and $\alpha'^{(i)}$, for some $0 \leq p < i \leq m-1$. Since β is periodic, every letter in $\mathcal{A}_\beta$ is singular for β. Then, without loss of generality, either $a^\infty = \beta^{(p)} \neq \beta^{(i)} = b^\infty$ or $a^\infty = \beta^{(p)} = \beta^{(i)}$.

Let τ be a maximal k-sized window over $\alpha'^{(p)}$.

Case IIa. Suppose that $a^\infty = \beta^{(p)} \neq \beta^{(i)} = b^\infty$. Then, b must appear in $\alpha'^{(p)}$ infinitely often (and likewise, a must appear in $\alpha'^{(i)}$ infinitely often) or otherwise there are at least $2k$ τ-words in $\alpha'^{(p)}$ with no b in them and $k-2$ τ-words of the form $b^j d u_j$ for $0 < j < k$, $d \neq b$, and $u_j \in \mathcal{A}^{k-j-1}$ in $\alpha^{(i)}$. Thus, $\#F_\alpha(m\tau) \geq 2k + k - 2$, which is a contradiction.

Since c must appear in some residue of α' aperiodically, then either there exists some q, $0 \leq q \leq m-1$ and $q \neq p, i$ such that $\beta^{(q)} = c^\infty$, or c appears in $\alpha'^{(p)}$ and $\alpha'^{(i)}$ infinitely often (since appearing in one and not the other would give at least $2k$ words with no c in one residue and $k-2$ words of the form $u_j c u'_j$ for $0 < j < k$ in the other residue).

If there exists some q, $0 \leq q \leq m-1$, and $q \neq p, i$ such that $\beta^{(q)} = c^\infty$, then there are $k-2$ τ-words of the form $a^j d u_j$ for $0 < j < k$, $d \neq b$, and $u_j \in \mathcal{A}^{k-j-1}$ in $\alpha^{(p)}$; $k-2$ τ-words of the form $b^j e u_j$ for $0 < j < k$, $e \neq b$, and $u_j \in \mathcal{A}^{k-j-1}$ in $\alpha^{(i)}$; and $k-2$ τ-words of the form $c^j f u_j$ for $0 < j < k$, $f \neq c$, and $u_j \in \mathcal{A}^{k-j-1}$ in $\alpha^{(q)}$. Thus, $\#F_\alpha(m\tau) \geq 3k-6$ for all $k > m$, which is a contradiction.

Thus, c appears in $\alpha'^{(p)}$ infinitely often. Since $\alpha'^{(p)}$ sees a, b, and c infinitely often, by Lemma 4.2, there is at least one letter in $\alpha'^{(p)}$ that appears periodically in $\alpha'^{(p)}$, which is a contradiction.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

Case IIb. Now, assume that $\beta^{(p)} = \beta^{(i)} = a^\infty$. Further, assume that there are no residues q for which $\alpha'^{(q)}$ is aperiodic and $\beta^{(q)} = d^\infty$ for $d \neq a$ or else we can reduce to Case IIa.

Thus, b and c must appear aperiodically in either $\alpha'^{(p)}$ or $\alpha'^{(i)}$ by Lemma 4.3.

If, without loss of generality, b only appears in $\alpha'^{(i)}$ and not in $\alpha'^{(p)}$, then there are at least $2k$ τ-words from $\alpha'^{(p)}$ that do not contain a b and $k-2$ τ-words from $\alpha'^{(i)}$ of the form $u_j b u'_j$ for $0 < j < k$, $u_j \in (\mathcal{A} \setminus \{b\})^j$ and $u'_j \in \mathcal{A}^{k-j-1}$.

Thus, b and c appear in $\alpha'^{(p)}$ infinitely often. Then, by Lemma 4.2, there is some letter in $\alpha'^{(p)}$ that appears periodically, which is a contradiction.

Therefore, by Cases I and II, $\#\mathcal{A}_A = 2$. □

THEOREM 4.5. *There exists some decomposition of α' into $1 < m_0 \leq m$ residues, where m is the singular decomposition cycle of α', such that exactly one residue of the decomposition is aperiodic over two letters and any letter not in $\mathcal{A}_A$ is singular for α'.*

Proof. Suppose that $\alpha'^{(p)}$, $0 \leq p \leq m-1$, is aperiodic.

Let $k > m$ and let τ be a maximal $(k-(m-1))$-sized window over $\alpha'^{(p)}$. Then, $\#F_{\alpha'^{(p)}}(\tau) \geq 2k - 2(m-1)$ by Theorem 2.8.

Define a k-sized window τ' over α' by $\tau' = m\tau \cup \{h_i : 1 \leq i \leq m-1\}$ for $h_1 > m\tau(k-(m-1)-1)$, $h_{i+1} > h_i$ for all $1 \leq i \leq m-2$, and $h_i \equiv i \mod m$ for all $1 \leq i \leq m-1$.

Define $x \in (\mathcal{A} \cup \{?\})^{\mathbb{N}}$ by $x_i = \alpha'_i$ if $i \in S$ and $x_i = ?$ if $i \in A$.

If x has least period m, then $C_i(\tau', \alpha') \cap C_j(\tau', \alpha') = \emptyset$ for all $0 \leq i < j \leq m-1$. Furthermore, $\#C_p(\tau', \alpha') \geq 2k - 2(m-1)$ since $C_p(\tau', \alpha')$ witnesses τ over $\alpha'^{(p)}$.

If a second residue $\alpha'^{(i)}$, $1 \leq i \leq m-1$, were also aperiodic, then $\#F_{\alpha'^{(i)}}(\tau) \geq k-(m-1)+1$ by Fact 2.10. Thus, $\#C_i(\tau', \alpha') \geq k-(m-1)+1$, and since $C_p(\tau', \alpha')$ and $C_i(\tau', \alpha')$ are disjoint, $\#F_{\alpha'}(\tau') \geq \#F_{\alpha'}(\tau') \geq 3k - 3(m-1) + 1$ for all $k > m$, which is a contradiction.

If x has least period $m_0 < m$, then m_0 divides m. Suppose that there is a second aperiodic residue, $\alpha'^{(i)}$. By Lemma 4.1, $i \in A$.

If $i \not\equiv p \mod m_0$, then $C_p(\tau', \alpha') \cap C_i(\tau', \alpha') = \emptyset$. Thus, since $\#C_p(\tau', \alpha') \geq 2k - 2(m-1)$ and by Fact 2.10, $\#C_i(\tau', \alpha') \geq k-(m-1)+1$, then $\#F_{\alpha'}(\tau') \geq 3k - 3(m-1) + 1$, which is a contradiction.

If $i \equiv p \mod m_0$ and for all $j \not\equiv 0 \mod m_0$, $\alpha'^{(j)}$ is periodic, and by Theorem 4.4, $\#\mathcal{A}_A = 2$, then α' can be decomposed with respect to m_0. We denote this decomposition $\{\alpha'^{[i]} : 0 \leq i \leq m_0 - 1\}$. Then, the only aperiodic residue is $\alpha'^{[p]}$ and all other residues either belong to S and thus are singular for α', or are periodic over the alphabet of $\alpha'^{[p]}$.

Thus, we have exactly one aperiodic residue over two letters.

Suppose $c \notin \mathcal{A}_{\alpha'^{(p)}}$. Then, c only appears in periodic residues. Thus, c appears periodically in α' and thus, must appear periodically in β. Hence, c is singular for β and must be singular for α'. □

By Theorem 4.5 and the discussion at the beginning of this section, the conclusion for Theorem 1.1 has been proved for when α is not recurrent, completing the proof of Theorem 1.1.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

5. *Strong pattern Sturmian sequences*

In this section, we prove Theorems 1.2 and 1.3. First, Lemma 5.1 and its immediate corollary introduce the relationship between the maximal pattern complexity of a sequence and the maximal pattern complexity of its aperiodic residue, which is unique by Theorem 1.1.

LEMMA 5.1. *If* $\liminf_{k\to\infty} p^*_\alpha(k) - 3k = -\infty$ *and* $\alpha^{(p)}$ *is the aperiodic residue from Theorem 1.1, then* $p^*_\alpha(k) \geq p^*_{\alpha^{(p)}}(k) + \ell - 2$.

Proof. Let $k \in \mathbb{N}$ and let τ be a maximal k-sized window over $\alpha^{(p)}$. Consider the window $m\tau$ over α. Then, $\#F_\alpha(m\tau) \geq \#F_{\alpha^{(p)}}(\tau) + \ell - 2$ since there is at least one word beginning with each letter of the alphabet. Thus, $p^*_\alpha(k) \geq \#F_\alpha(m\tau) \geq p^*_{\alpha^{(p)}}(k) + \ell - 2$. □

An immediate corollary of Lemma 5.1 is as follows.

COROLLARY 5.2. *If* $p^*_\alpha(k) = 2k + C$ *for some* $C \geq \ell - 2$ *and* $\alpha^{(p)}$ *is the only aperiodic residue,* $0 \leq p \leq m - 1$, *then* $p^*_{\alpha^{(p)}}(k) \leq 2k + D$ *for some* $0 \leq D \leq C - (\ell - 2)$ *and all* $k \in \mathbb{N}$.

Lemma 5.3 sets the foundation for proving Theorem 1.2 by showing that a non-trivial prefix of a non-recurrent sequence, as discussed in the beginning of §4, cannot occur when the maximal pattern complexity is restricted to $2k + \ell - 2$.

LEMMA 5.3. *If* $p^*_\alpha(k) = 2k + \ell - 2$ *for all* $k \in \mathbb{N}$, *then the aperiodic residue from Theorem 1.1 is over two letters, a and b, and for any* $c \in \mathcal{A}$ *that is not equal to a or b, c is singular.*

Proof. If α is recurrent, this is clearly true by Theorem 3.1 and Proposition 3.2. Therefore, we will assume that α is not recurrent. By the discussion at the beginning of §4, there is a word w such that $\alpha = w\alpha'$, all residues of α' are aperiodic or periodic, and since $p^*_{\alpha'}(k) \leq p^*_\alpha(k)$, by Theorem 4.5, α' is a sequence with exactly one aperiodic residue over two letters, a and b, and for any $c \in \mathcal{A}$, $c \neq a, b$, c is singular. Thus, it suffices to show that w is trivial.

Assume, toward a contradiction, that w is non-trivial.

Define $x \in (\mathcal{A}_\alpha \cup \{?\})^{\mathbb{N}_0}$ by $x_i = ?$ if $\mathcal{A}_{\alpha'^{(i)}} \subseteq \{a, b\}$ and $x_i = \alpha'_i$ if $\mathcal{A}_{\alpha'^{(i)}} \not\subseteq \{a, b\}$. Without loss of generality, we can assume x has least period m, or else α could be decomposed with respect to $m_0 < m$, and $\alpha = w'\alpha''$, where w is still non-trivial and α'' has the same structure as α'.

Let $\alpha'^{(p)}$, $0 \leq p \leq m - 1$, be the only aperiodic residue of α'.

Since w is non-trivial, there is some residue i such that either there exists some $c \in \mathcal{A}_{\alpha^{(i)}}$ such that $c \notin \mathcal{A}_{\alpha'^{(i)}}$ and $\mathcal{A}_{\alpha^{(i)}} \neq \{a, b\}$ or $\mathcal{A}_{\alpha^{(i)}} = \{a, b\}$, $\alpha'^{(i)}$ is periodic (potentially constant), and for any $t \in \mathbb{N}$, $\sigma^t(\alpha^{(i)}) \neq \alpha^{(i)}$.

Consider the first case, where there exists some $c \in \mathcal{A}_{\alpha^{(i)}}$ such that $c \notin \mathcal{A}_{\alpha'^{(i)}}$ and $\mathcal{A}_{\alpha^{(i)}} \neq \{a, b\}$.

Let $k \geq 2$ and let τ be a maximal k-sized window over $\alpha'^{(p)}$. Then, $F_{\alpha'}(m\tau) = F_{\alpha'^{(p)}}(\tau) \cup \{d^k : d \in \mathcal{A}_\alpha \text{ and } d \neq a, b\}$, and so $\#F_{\alpha'}(m\tau) \geq 2k + \ell - 2$. We also note that $F_{\alpha'^{(p)}}(\tau)$ only contains words over the letters a and b.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

However, since there is some residue i for which there exists some $c \in \mathcal{A}_{\alpha^{(i)}}$ such that $c \notin \mathcal{A}_{\alpha'^{(i)}}$, then either the $m\tau$-word cd^{k-1} for $c \notin \{a, b\}$, cd^{k-1} for $d \notin \{a, b\}$, or cu for $c \notin \{a, b\}$ and $u \in F_{\alpha'^{(p)}}(\tau \setminus \{\tau(k-1)\})$ must also appear in $F_\alpha(m\tau)$, and so $\#F_\alpha(m\tau) \geq 2k + \ell - 1$, which is a contradiction.

Now, consider the second case, where $\mathcal{A}_{\alpha^{(i)}} = \{a, b\}$, $\alpha'^{(i)}$ is periodic (potentially constant), and for any $t \in \mathbb{N}$, $\sigma^t(\alpha^{(i)}) \neq \alpha^{(i)}$. We will further assume that the first case cannot occur, so constant residues in α' over letters other than a and b must also be constant in α.

Let q be the least period of $\alpha'^{(i)}$. Since m is the least period of x, there exists some j, $0 \leq j \leq m-1$, such that $\mathcal{A}_{\alpha^{(i+j \mod m)}} \cap \mathcal{A}_{\alpha^{(p+j \mod m)}} = \emptyset$. Without loss of generality, suppose that $p + j \not\equiv i \mod m$ (otherwise, use the distance $-j \mod m$ instead).

Define a q-sized window τ by $\tau = \{0, m, 2m, \ldots, (q-1)m\}$. Define an immediate extension τ' of τ by $\tau' = \tau \cup \{h\}$ for some $h > (q-1)m$ and $h \equiv j \mod m$. Then, $C_i(\tau', \alpha') \cap C_p(\tau', \alpha') = \emptyset$ and $\{\sigma^t(\alpha'^{(i)})_0 \sigma^t(\alpha'^{(i)})_1 \ldots \sigma^t(\alpha'^{(i)})_{q-1} : 0 \leq t \leq q-1\} \cup \{\alpha_0^{(i)}\alpha_1^{(i)} \ldots \alpha_{q-1}^{(i)}\} \subset C_i(\tau, \alpha)$, so $\#C_i(\tau, \alpha) \geq q + 1$.

If $i + j \equiv p \mod m$, then $\#C_i(\tau', \alpha) \geq q + 2$ since at least one word from $C_i(\tau, \alpha)$ must have at least two extensions in $C_i(\tau', \alpha)$ or else $\alpha^{(p)}$ would be eventually periodic. Furthermore, $\#C_p(\tau', \alpha) \geq q + 1$ by Fact 2.10 since $\alpha^{(p)}$ is aperiodic. Since there are at least $\ell - 2$ residues that are constant over letters not equal to a or b, there are at least $\ell - 2$ more τ'-words. Thus, $\#F_\alpha(\tau') \geq 2q + 2 + \ell - 1$, which is a contradiction.

If $i + j \not\equiv p \mod m$, then $\#C_i(\tau', \alpha) \geq q + 1$, $\#C_p(\tau', \alpha) \geq q + 1$ by Fact 2.10. Since $C_i(\tau, \alpha) \cap C_p(\tau, \alpha) = \emptyset$, either $\alpha^{(p-j \mod m)} = c^\infty$ or $\alpha^{(i-j \mod m)} = c^\infty$ for some $c \notin \{a, b\}$. Thus, either $\#C_{i-j \mod m}(\tau', \alpha) \geq 2$ or $\#C_{p-j \mod m}(\tau', \alpha) \geq 2$. Then, there are at least $\ell - 3$ τ'-words over the residues that are constant over letters not equal to a, b, or c. Thus, $\#F_\alpha(\tau') \geq 2q + 2 + \ell - 1$, which is a contradiction.

Thus, w must be trivial. □

We now prove Theorem 1.2.

Proof of Theorem 1.2. Suppose that α is an aperiodic sequence such that $\#\mathcal{A}_\alpha = \ell \geq 3$ and $p_\alpha^*(k) = 2k + \ell - 2$ for all $k \in \mathbb{N}$.

By Lemma 5.3, exactly one residue is aperiodic over a two-letter alphabet $\{a, b\}$. Let $\alpha^{(p)}$, $0 \leq p \leq m-1$, be the aperiodic residue. Also by Lemma 5.3, for all $c \in \mathcal{A}$ such that $c \notin \mathcal{A}_{\alpha^{(p)}}$, c is singular for α.

By Lemma 5.1, since $2k + \ell - 2 = p_\alpha^*(k) \geq p_{\alpha^{(p)}}^*(k) + \ell - 2$, then $p_{\alpha^{(p)}}^*(k) \leq 2k$. Since $\alpha^{(p)}$ is aperiodic, by Theorem 2.8, $p_{\alpha^{(p)}}^*(k) \geq 2k$, and so the aperiodic residue must have maximal pattern complexity $2k$ and is a pattern Sturmian over two letters.

Since each letter not witnessed by $\alpha^{(p)}$ is singular, any residue on letters other than those in $\mathcal{A}_{\alpha^{(p)}}$ are constant. Thus, it suffices to show that there is no residue $\alpha^{(j)}$, $j \neq p$, which is non-constant periodic over $\mathcal{A}_{\alpha^{(p)}}$.

Suppose that there is some residue j, $0 \leq j \leq m-1$ and $j \neq p$, such that $\alpha^{(j)}$ is non-constant periodic over $\mathcal{A}_{\alpha^{(p)}} = \{a, b\}$. Let $k > m$, τ be a maximal $(k - (m-1))$-sized window over $\alpha^{(p)}$, and define a k-sized window τ' over α by $\tau' = m\tau \cup \{h_i : 1 \leq i \leq m-1\}$ for $h_1 > m\tau(k - (m-1) - 1)$, $h_{i+1} > h_i$ for all

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

$1 \le i \le m-2$, and $h_i \equiv i \mod m$ for all $1 \le i \le m-1$. Then, $\#C_p(\tau', \alpha) \ge 2k - 2(m-1)$. Since $\alpha^{(j)}$ is non-constant and $\alpha^{(p)}$ is aperiodic, in every $C_i(\tau', \alpha)$, there are at least three distinct words classified by the letters appearing in the parts of the window centered over $\alpha^{(p)}$ and $\alpha^{(j)}$: two distinct words with an a or b in the $\alpha^{(j)}$ slot, and a third word since if the letter appearing in the $\alpha^{(j)}$ slot always forced the letter appearing in the $\alpha^{(p)}$ slot, then $\alpha^{(p)}$ would be periodic (either constant, equal to $\alpha^{(j)}$, or the bit flip of $\alpha^{(j)}$). Therefore, $\#F_\alpha(\tau) \ge 2k - 2(m-1) + 3(m-1) = 2k + m - 1$. Furthermore, $m \ge \ell$ since $\alpha^{(p)}$ and $\alpha^{(j)}$ only use the letters a and b, and there are at least $\ell - 2$ other constant residues over the letters not equal to a or b. Therefore, $\#F_\alpha(\tau) \ge 2k + \ell - 1$, which is a contradiction. □

The next part of this section is dedicated to proving Theorem 1.3.

Proof of Theorem 1.3. Suppose α is aperiodic, $\#\mathcal{A}_\alpha = \ell \ge 3$, and there exists some $m \ge 2$ such that the decomposition of α with respect to m has one residue that is a pattern Sturmian over two letters and all other residues are constant.

Let $\alpha^{(p)}$, $0 \le p \le m-1$, be the aperiodic pattern Sturmian over two letters.

Let $k \in \mathbb{N}$ and let τ be any k-sized window. Define $\tau_i = \{\tau(j) : j \equiv i \mod m\}$ for all $0 \le i \le m-1$. Then, $\tau = \tau_0 \sqcup \tau_1 \sqcup \ldots \sqcup \tau_{m-1}$.

Suppose that $I = \{i : 0 \le i \le m-1 \text{ and } \tau_i \ne \emptyset\}$.

Since any word in $C_i(\tau, \alpha)$ is completely determined by the letters of $\alpha^{(p)}$ which appear in it, $\#C_i(\tau, \alpha) \le 2(\#\tau_{p-i \mod m})$.

Furthermore, $\#(\bigcup_{i \notin I} C_{p-i \mod m}(\tau, \alpha)) \le m - 1 - \#I$ since for all $i \notin I$, $C_{p-i \mod m}(\tau, \alpha)$ is a set of words that only see constant sequences, so each $C_{p-i \mod m}(\tau, \alpha)$ can contribute at most one word.

Therefore, for all k-sized windows τ,

$$\#F_\alpha(\tau) \le \sum_{i \in I} 2(\#\tau_i) + \#\left(\bigcup_{i \notin I} C_i(\tau, \alpha)\right) \le 2k + m - 1 - \#I \le 2k + m - 2.$$

Thus, $p^*_\alpha(k) \le 2k + m - 2$. By an application of Theorem 2.9, $p^*_\alpha(k+1) \ge p^*_\alpha(k) + 2$ for all $k \in \mathbb{N}$, and so for sufficiently large k, $p^*_\alpha(k) = 2k + C$ for some $\ell - 2 \le C \le m - 2$. □

6. *Examples*

This section introduces some relevant examples. Example 6.1 demonstrates the existence of strong pattern Sturmian sequences over any finite alphabet.

Example 6.1. We define a sequence $\alpha \in \{0, 1, \ldots, \ell - 1\}^{\mathbb{N}_0}$ that has maximal pattern complexity $p^*_\alpha(k) = 2k + \ell - 2$.

Let $x \in \{0, 1\}^{\mathbb{N}_0}$ be a pattern Sturmian sequence over two letters. Define $\alpha \in \{0, 1, \ldots, \ell - 1\}^{\mathbb{N}_0}$ by

$$\alpha_{(\ell-1)i+j} = \begin{cases} x_i & \text{if } j \equiv 0 \mod \ell - 1, \\ j + 1 & \text{if } j \not\equiv 0 \mod \ell - 1. \end{cases}$$

The singular decomposition cycle of α is $\ell - 1$.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

Let $k \in \mathbb{N}$ and let τ be any k-sized window. Define $\tau_i = \{\tau(j) : j \equiv i \mod m\}$ for all $0 \le i \le m-1$. Then, $\tau = \tau_0 \sqcup \tau_1 \sqcup \cdots \sqcup \tau_{m-1}$.

Suppose that $I = \{i : 0 \le i \le \ell - 2 \text{ and } \tau_i \neq \emptyset\}$.

Since any two residues have disjoint alphabets, for all $0 \le i < j \le \ell - 2$, $C_i(\tau, \alpha) \cap C_j(\tau, \alpha) = \emptyset$.

Furthermore, $C_i(\tau, \alpha) \le \max\{2(\#\tau_{\ell-1-i}), 1\}$ and so $\#F_\alpha(\tau) = \sum_{i=0}^{\ell-2} \#C_i(\tau, \alpha) \le 2k + (\ell - 1 - \#I) \le 2k + \ell - 2$ since $\#I \ge 1$.

Since this is true for any k-sized window τ, we have that $p^*_\alpha(k) \le 2k + \ell - 2$ and since by Theorem 2.11, $p^*_\alpha(k) \ge 2k + \ell - 2$, for this sequence, $p^*_\alpha(k) = 2k + \ell - 2$.

Example 6.2 is a sequence that has the structure outlined in Theorem 1.3, but has maximal pattern complexity strictly greater than $2k + \ell - 2$ demonstrating that the converse of Theorem 1.2 is not true.

Example 6.2. We will define a sequence $\alpha \in \{0, 1, 2, 3\}^{\mathbb{N}_0}$ that is a pattern Sturmian sequence zippered with constant sequences, but has maximal pattern complexity $p^*_\alpha(k) = 2k + C$ for $C > \ell - 2$.

Let $x \in \{0, 1\}^{\mathbb{N}_0}$ be a pattern Sturmian over two letters. Define α by

$$\alpha_{5i+j} = \begin{cases} x_i & \text{if } j \equiv 0 \mod 5, \\ 2 & \text{if } j \equiv 1, 4 \mod 5, \\ 3 & \text{if } j \equiv 2, 3 \mod 5. \end{cases}$$

Let τ be a maximal window for x of size $k - 1$. Define τ' by $\tau'(i) = m\tau(i)$ for all $0 \le i \le k - 2$ and $\tau'(k-1) = h$ for some $h \equiv 1 \mod 5$ and $h > m\tau(k-2)$.

Then, $\#C_0(\tau', \alpha) = 2(k-1)$ since $\alpha^{(0)}$ is a pattern Sturmian sequence over two letters and τ is a maximal window for $\alpha^{(0)}$. Furthermore, $\#C_4(\tau, \alpha) = 2$ since $C_4(\tau, \alpha) = \{2^{k-1}0, 2^{k-1}1\}$, and $\#C_1(\tau, \alpha) = \#C_2(\tau, \alpha) = \#C_3(\tau, \alpha) = 1$ since $C_1(\tau, \alpha) = \{2^{k-1}3\}$, $C_2(\tau, \alpha) = \{3^k\}$, and $C_3(\tau, \alpha) = \{3^{k-1}2\}$. Clearly, $\{C_i(\tau', \alpha) : 0 \le i \le 4\}$ are pairwise disjoint. Thus, $\#F_\alpha(\tau') = 2k + 3 > 2k + (4 - 2) = 2k + \ell - 2$ for $\ell = 4$.

Example 6.3 is a uniformly recurrent sequence with two aperiodic residues that has complexity $4k$, showing that the bound in Theorem 1.1 cannot be substantially improved below $4k$.

Example 6.3. We create a sequence that has complexity $4k$ and is two recurrent pattern Sturmian sequences zippered together with a constant sequence.

Let $x \in \{0, 1\}^{\mathbb{N}_0}$ be the Fibonacci substitution defined by $0 \to 01$ and $1 \to 0$, a Sturmian and therefore pattern Sturmian sequence, which is also uniformly recurrent.

Define $\alpha \in \{0, 1, 2\}^{\mathbb{N}_0}$ by

$$\alpha_{3i+j} = \begin{cases} x_i & \text{if } j \equiv 0, 1 \mod 3, \\ 2 & \text{if } j \equiv 2 \mod 3. \end{cases}$$

First, we must show that α is uniformly recurrent. Let $L > 0$ and let $R_L(\alpha) = \{M \in \mathbb{N} : \alpha_0 \ldots \alpha_{L-1} = \alpha_M \ldots \alpha_{M+L-1}\}$. Let $L' = \lceil L/3 \rceil$ and define $R'_L(x) =$

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

$\{M \in \mathbb{N} : x_0 \dots x_{L'-1} = x_M \dots x_{M+L'-1}\}$. Since x is uniformly recurrent, $R_{L'}(x) = \{M_1, M_2, \dots\}$ is syndetic and for any M_i, $x_0x_1 \dots x_{L'-1} = x_{M_i}x_{M_i+1} \dots x_{M_i+L'-1}$. Thus, $x_0x_02x_1x_12 \dots x_{L'-1}x_{L'-1}2 = x_{M_i}x_{M_i}2x_{M_i+1}x_{M_i+1}2 \dots x_{M_i+L'-1}x_{M_i+L'-1}2$, and so $R_L(\alpha) = \{3M_1, 3M_2, \dots\}$ is syndetic and thus, α is a uniformly recurrent sequence.

Now, we show that α has maximal pattern complexity $p^*_\alpha(k) = 4k$ for all $k \geq 2$.

First, we show that for any k-sized window τ over α, $p^*_\alpha(k) \leq 4k$.

Let $k \in \mathbb{N}$ and let τ be any k-sized window over α. If for all $0 \leq i \leq k-1$, $\tau(i) \equiv 0 \mod 3$ (i.e. τ only sees one residue at a time), then $C_0(\tau, \alpha) = C_1(\tau, \alpha)$ since $\alpha^{(0)} = \alpha^{(1)} = x$, and $\#C_0(\tau, \alpha) \leq 2k$ since x is a pattern Sturmian sequence. Furthermore, $C_2(\tau, \alpha) = \{2^k\}$. Thus, $\#F_\alpha(\tau) \leq 2k + 1$.

Suppose now that $\tau = \tau_1 \sqcup \tau_2 \sqcup \tau_3$ for $\tau_i = \{\tau(j) : \tau(j) \equiv i \mod 3\}$ and there exists $i \neq j$ such that $\#\tau_i, \#\tau_j > 0$. Then, for all $0 \leq i < j \leq 2$, $C_i(\tau, \alpha) \cap C_j(\tau, \alpha) = \emptyset$ since the position of the 2s uniquely determines to which set $C_i(\tau, \alpha)$ a τ-word belongs.

Then, $\#C_0(\tau', \alpha) \leq 2(\#\tau_0) + 2(\#\tau_1)$, $\#C_1 \leq 2(\#\tau_0) + 2(\#\tau_2)$, and $\#C_2 \leq 2(\#\tau_1) + 2(\#\tau_2)$. Therefore,

$$\#F_\alpha(\tau) \leq 2(\#\tau_0) + 2(\#\tau_1) + 2(\#\tau_0) + 2(\#\tau_2) + 2(\#\tau_1) + 2(\#\tau_2) = 4k.$$

Thus, for all $k \in \mathbb{N}$ and any k-sized window τ, $\#F_\alpha(\tau) \leq 4k$, and so $p^*_\alpha(k) \leq 4k$ for all $k \in \mathbb{N}$.

Now, we show that for all $k \geq 2$, there exists some window τ' such that $\#F_\alpha(\tau') = 4k$.

Let τ be a maximal k-sized window for x such that the immediate restriction of τ is a maximal $(k-1)$-sized window. Define τ' by $\tau'(i) = 3\tau(i)$ for all $0 \leq i \leq k-2$ and $\tau'(k-1) = 3\tau(k-1) + 1$.

For all $0 \leq i < j \leq 2$, $C_i(\tau, \alpha) \cap C_j(\tau, \alpha) = \emptyset$ since words in $C_0(\tau', \alpha)$ are of the form w_0 for $w_0 \in \{0, 1\}^k$, words in C_1 are of the form w_12 for $w_1 \in \{0, 1\}^{k-1}$, and words in C_2 are of the form $2^{k-1}w_3$ for $w_3 \in \{0, 1\}$.

Then, $\#C_0(\tau', \alpha) = 2k$ since $C_0(\tau', \alpha) = F_x(\tau)$, $\#C_1(\tau', \alpha) = 2k - 2$ since $C_1(\tau', \alpha) = \{w_12 : w_1 \in F_x(\tau \setminus \{\tau(k-1)\})\}$, and $\#C_2(\tau', \alpha) = 2$ since $C_2(\tau', \alpha) = \{2^{k-1}0, 2^{k-1}1\}$.

Thus, $\#F_\alpha(\tau') = 4k$ and so $p^*_\alpha(k) = 4k$ for all $k \in \mathbb{N}$.

Acknowledgements. C.S. would like to express her gratitude to her advisor Prof. Ronnie Pavlov for suggesting this question and for the numerous helpful comments, suggestions, and conversations during the formation of this paper.

References

[Fog02] N. Pytheas Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics (Lecture Notes in Mathematics, 1794)*. Ed. V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel. Springer-Verlag, Berlin, 2002.

[KR06] T. Kamae and H. Rao. Maximal pattern complexity of words over l letters. *European J. Combin.* **27**(1) (2006), 125–137.

[KZ02a] T. Kamae and L. Zamboni. Maximal pattern complexity for discrete systems. *Ergod. Th. & Dynam. Sys.* **22**(4) (2002), 1201–1214.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press

[KZ02b] T. Kamae and L. Zamboni. Sequence entropy and the maximal pattern complexity of infinite words. *Ergod. Th. & Dynam. Sys.* **22**(4) (2002), 1191–1199.
[MH38] M. Morse and G. A. Hedlund. Symbolic dynamics. *Amer. J. Math.* **60**(4) (1938), 815–866.
[MH40] M. Morse and G. A. Hedlund. Symbolic dynamics II. Sturmian trajectories. *Amer. J. Math.* **62**(1) (1940), 1–42.

https://doi.org/10.1017/etds.2026.10303 Published online by Cambridge University Press