On the 2-binomial complexity of the generalized Thue-Morse words

Xiao-Tao Lü^a, Jin Chen^a, Zhi-Xiong Wen^b, Wen Wu^{c,*}

^aCollege of Science, Huazhong Agricultural University, Wuhan 430070, China.

^bSchool of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, 430074,

China.

^cSchool of Mathematics, South China University of Technology, Guangzhou, 510641, China

Abstract

In this paper, we study the 2-binomial complexity $b_{\mathbf{t}_m,2}(n)$ of the generalized Thue-Morse words \mathbf{t}_m for every integer $m \geq 3$. We obtain the exact value of $b_{\mathbf{t}_m,2}(n)$ for every integer $n \geq m^2$. As a consequence, $b_{\mathbf{t}_m,2}(n)$ is ultimately periodic with period m^2 . This result partially answers a question of M. Lejeune, J. Leroy and M. Rigo [Computing the k-binomial complexity of the Thue-Morse word, J. Comb. Theory Ser. A, **176** (2020) 105284].

Keywords: generalized Thue-Morse word, k-binomial equivalence, k-binomial complexity

1. Introduction

Abelian equivalence of words has been a subject of great interest for quite a long time. Given a finite non-empty set \mathcal{A} , let \mathcal{A}^* and $\mathcal{A}^{\mathbb{N}}$ denote the set of finite words and the set of infinite words over \mathcal{A} respectively. Two words $u, v \in \mathcal{A}^*$ are *abelian equivalent*, denoted by $u \sim_{ab} v$, if $|u|_a = |v|_a$ for every $a \in \mathcal{A}$ where $|u|_a$ denotes the number of occurrences of the letter a in u. The notion has been studied in the relation of abelian complexity of infinite words [4, 14, 15, 23], abelian repetitions and avoidance [3, 10, 18], and other topics [6, 17, 21]; see also [20] and references therein.

As a generalization of abelian equivalence, Rigo and Salimov [25] introduced the notion of kbinomial equivalence based on binomial coefficients of words. The binomial coefficient $\binom{u}{v}$ of two words u and v is the number of times that v occurs as a (not necessarily contiguous) subsequence of u. Binomial coefficients of finite words have been successfully applied in several fields: padic topology [2], non-commutative extension of Mahler's theorem on interpolation series [19], formal language theory [8], Parikh matrices and a generalization of Sierpiński's triangle [13]. Many classical questions in combinatorics on words can be considered in the binomial context. Avoiding binomial squares and cubes was considered in [22]. The problem of testing whether two words are k-binomially equivalent was discussed in [7]. Let $k \in \mathbb{Z}^+ \cup \{+\infty\}$. Two words uand v are k-binomially equivalent, denoted by $u \sim_k v$, if $\binom{u}{x} = \binom{v}{x}$ for all words x of length at most k. Note that $u \sim_{+\infty} v$ if and only if u = v, while \sim_1 corresponds to the usual notion of abelian equivalence \sim_{ab} . Thus one can regard the notion of k-binomial equivalence as gradually bridging the gap between abelian equivalence (k = 1) and equality $(k = +\infty)$. An independent generalization of abelian equivalence is k-abelian equivalence where one counts factors of length at most k [9]; for more details, see [24].

^{*}Corresponding author.

Email addresses: xiaotaoLv@mail.hzau.edu.cn (Xiao-Tao Lü), cj@mail.hzau.edu.cn (Jin Chen), zhi-xiong.wen@hust.edu.cn (Zhi-Xiong Wen), wuwen@scut.edu.cn (Wen Wu)

Given an infinite word $\mathbf{w} = w(0)w(1)w(2)\cdots \in \mathcal{A}^{\mathbb{N}}$, for every positive integer n, let $\mathcal{F}_{\mathbf{w}}(n)$ denote the set of all factors of \mathbf{w} of length n. That is, $\mathcal{F}_{\mathbf{w}}(n) = \{w(i)w(i+1)\cdots w(i+n-1) \mid i \geq 0\}$. Set $\rho_{\mathbf{w}}(n) = \sharp(\mathcal{F}_{\mathbf{w}}(n))$. The function $\rho_{\mathbf{w}} \colon \mathbb{Z}^+ \to \mathbb{Z}^+$ is called the factor complexity function of \mathbf{w} . A fundamental result due to Hedlund and Morse [16] states that an infinite word \mathbf{w} is ultimately periodic if and only if $\rho_{\mathbf{w}}(n) \leq n$ for some $n \geq 1$. Words of factor complexity $\rho_{\mathbf{w}}(n) = n+1$ are called Sturmian words. Analogously, for each $k \in \mathbb{Z}^+ \cup \{+\infty\}$, the *k*-binomial complexity of \mathbf{w} is define as $b_{\mathbf{w},k}(n) = \sharp(\mathcal{F}_{\mathbf{w}}(n)/\sim_k)$. The function $b_{\mathbf{w},k}(n) \colon \mathbb{Z}^+ \to \mathbb{Z}^+$ counts the number of *k*-binomial equivalence classes of factors of length *n* occurring in \mathbf{w} . In the case $k = +\infty$, it holds that $b_{\mathbf{w},+\infty}(n) = \rho_{\mathbf{w}}(n)$, while if $k = 1, b_{\mathbf{w},1}(n)$, denoted by $\rho_{\mathbf{w}}^{ab}(n)$, corresponds to the usual abelian complexity of \mathbf{w} .

Abelian complexity is now a widely studied property of infinite words that has been examined for the first time by Coven and Hedlund in [5], where they have revealed that it could serve as an alternative way to characterize periodic words and Sturmian words. Coven and Hedlund showed that an infinite word **w** is periodic if and only if its abelian complexity satisfies $\rho_{\mathbf{w}}^{ab}(n) = 1$ for all large enough n, and they proved that an aperiodic binary infinite word **s** is Sturmian if and only if $\rho_{\mathbf{s}}^{ab}(n) = 2$ for every integer $n \geq 1$. The notion "abelian complexity" itself comes from the paper [23] which initiated a general study of the abelian complexity of infinite words over finite alphabets. The abelian complexity functions of some notable words have been determined, for example, the Thue–Morse word [23], the paperfolding word [15], the Rudin-Shapiro word [14] and the generalized Thue-Morse words [4].

However nontrivial infinite words with a closed form of the k-binomial complexity $b_{\mathbf{w},k}(n)$ are very rare. There are only a few such examples to date.

• Let **t** be the Thue-Morse word. Let k be a positive integer. Lejeune, Leroy and Rigo [11] proved that for all $n \leq 2^k - 1$, $b_{\mathbf{t},k}(n) = \rho_{\mathbf{t}}(n)$ and for all $n \geq 2^k$,

$$b_{\mathbf{t},k}(n) = \begin{cases} 3 \cdot 2^k - 3, \text{ if } n \equiv 0 \pmod{2^k}; \\ 3 \cdot 2^k - 4, \text{ otherwise.} \end{cases}$$

• For a Sturmian word s, Rigo and Salimov [25] showed that for all $k \ge 2$,

$$b_{\mathbf{s},k}(n) = \rho_{\mathbf{s}}(n) = n+1 \ (n \ge 1).$$

- Let $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ be a morphism. If $\varphi(a) \sim_{ab} \varphi(b)$ for all $a, b \in \mathcal{A}$, then φ is said to be *Parikh-constant*. Rigo and Salimov [25] proved that there is a constant $C_{\mathbf{x},k} > 0$ such that $b_{\mathbf{x},k}(n) \leq C_{\mathbf{x},k}$ for all $n \geq 1$, where \mathbf{x} is a fixed point of a Parikh-constant morphism.
- For the Tribonacci word **T**, Lejeune, Rigo and Rosenfeld [12] proved that for all $k \geq 2$,

$$b_{\mathbf{T},k}(n) = \rho_{\mathbf{T}}(n) = 2n+1 \ (n \ge 1).$$

For every $k \geq 2$, finding explicit value of the k-binomial complexity $b_{\mathbf{w},k}(n)$ for a given infinite word \mathbf{w} is a difficult task, particularly in case of words defined over alphabets consisting of more than two letters. In [11, Section 8], Lejeune, Leroy and Rigo asked if it is possible to compute the exact value of $b_{\mathbf{x},k}(n)$ for the fixed point \mathbf{x} of any Parikh-constant morphism such as the generalized Thue-Morse word \mathbf{t}_m with $m \geq 3$. Let σ_m be a morphism over the alphabet $\Sigma_m := \{0, 1, \ldots, m-1\}$ defined as $0 \mapsto 01 \cdots (m-1), 1 \mapsto 12 \cdots (m-1)0, \ldots, m-1 \mapsto (m-1)0 \cdots (m-2)$. The generalized Thue-Morse word

$$\mathbf{t}_m := t_m(0)t_m(1)t_m(2)\cdots = 01\cdots$$

is the fixed point of the morphism σ_m beginning with 0, i.e., $\mathbf{t}_m = \sigma_m^{\infty}(0)$.

The abelian complexity (or 1-binomial complexity) $b_{\mathbf{t}_m,1}(n)$ was given in [4].

Theorem 1 ([4]). For all integer $n \ge m$, if $n \equiv r \pmod{m}$, then

$$b_{\mathbf{t}_m,1}(n) = \rho_{\mathbf{t}_m}^{ab}(n) = \begin{cases} \frac{1}{4}m(m^2 - 1) + 1, & \text{if } m \text{ is odd and } r = 0; \\ \frac{1}{4}m(m - 1)^2 + m, & \text{if } m \text{ is odd and } r \neq 0; \\ \frac{1}{4}m^3 + 1, & \text{if } m \text{ is even and } r = 0; \\ \frac{1}{4}m(m - 1)^2 + \frac{5}{4}m, & \text{if } m \text{ is even and } r \neq 0 \text{ is even} \\ \frac{1}{4}m^2(m - 2) + m, & \text{if } m \text{ is even and } r \neq 0 \text{ is odd.} \end{cases}$$

In this paper, we fully characterize the 2-binomial complexity $b_{\mathbf{t}_m,2}(n)$ with $m \geq 3$. Firstly we give a sufficient and necessary condition of the 2-binomial equivalence of two factors u and vof the word \mathbf{t}_m . Let S and \mathcal{P} denote the set of all the possible suffixes and prefixes defined as

$$\mathcal{S} := \{ w \in \Sigma_m^* \mid w \text{ is a suffix of } \sigma_m(a) \text{ for some } a \in \Sigma_m \text{ and } |w| < m \}, \\ \mathcal{P} := \{ w \in \Sigma_m^* \mid w \text{ is a prefix of } \sigma_m(a) \text{ for some } a \in \Sigma_m \text{ and } |w| < m \}.$$

Theorem 2. Let $u = \alpha \sigma_m(u')\beta$ and $v = \alpha' \sigma_m(v')\beta'$ be two factors of the generalized Thue-Morse word \mathbf{t}_m , where $(\alpha, \beta), (\alpha', \beta') \in S \times \mathcal{P}$ and $\min\{|u'|, |v'|\} \geq 3$. Then $u \sim_2 v$ if and only if $\alpha = \alpha', \beta = \beta'$ and $u' \sim_1 v'$.

We remark that there are infinite words which are fixed points of some Parikh-constant morphisms and do not satisfy the property of Theorem 2. For example, let $\mathbf{x} = \sigma^{\infty}(0)$ where σ is the morphism over $\{0, 1, 2\}$ defined as $0 \mapsto 012, 1 \mapsto 210, 2 \mapsto 120$. Let $u = \sigma(10122)21$ and $v = \sigma(22101)12$. Then u and v are two factors of \mathbf{x} and $u \sim_2 v$ with $\alpha = \alpha' = \varepsilon$, $\beta = 12$ and $\beta' = 21$. It is natural to ask that what kind of words share a property similar to Theorem 2.

Secondly, we obtain the exact value of the 2-binomial complexity $b_{\mathbf{t}_m,2}(n)$ for every $n \ge m^2$.

Theorem 3. For every $n \ge m^2$ with $m \ge 3$, we have

$$b_{\mathbf{t}_m,2}(n) = \begin{cases} b_{\mathbf{t}_m,1}(n/m) + m(m-1)(m(m-1)+1), & \text{if } n \equiv 0 \pmod{m}; \\ m^4 - 2m^3 + 2m^2, & \text{otherwise,} \end{cases}$$

where the abelian complexity $b_{t_m,1}(\cdot)$ is given in Theorem 1.

It follows from Theorem 1 and Theorem 3 that the 2-binomial complexity $b_{t_m,2}(n)$ is ultimately periodic.

Corollary 1. The 2-binomial complexity $b_{\mathbf{t}_m,2}(n)$ of the generalized Thue-Morse word \mathbf{t}_m is ultimately periodic with period m^2 .

As the abelian complexity (or 1-binomial complexity) $b_{\mathbf{t}_m,1}(n)$ and the 2-binomial complexity $b_{\mathbf{t}_m,2}(n)$ of the word \mathbf{t}_m are ultimately periodic with period m and m^2 respectively, our numerical result suggests the following conjecture.

Conjecture 1. For every $k \ge 3$, the k-binomial complexity $b_{t_m,k}(n)$ of the generalized Thue-Morse word is ultimately periodic with period m^k .

This paper is organized as follows. In Section 2, we state some basic definitions and notations. In Section 3, we prove Theorem 2. In the last section, we compute the exact value of $b_{\mathbf{t}_m,2}(n)$ for every $n \ge m^2$.

2. Preliminaries

2.1. Finite and infinite words

An alphabet \mathcal{A} is a finite and non-empty set whose elements are called *letters*. Any concatenation of letters from \mathcal{A} is called a word. The concatenation of two words $u = u(0)u(1)\cdots u(m)$ and $v = v(0)v(1)\cdots v(n)$ is the word $uv = u(0)u(1)\cdots u(m)v(0)v(1)\cdots v(n)$. The set of all finite words over \mathcal{A} including the *empty word* ε is denoted by \mathcal{A}^* . An infinite sequence of letters from \mathcal{A} is called an infinite word and the set of all infinite words over \mathcal{A} is denoted by $\mathcal{A}^{\mathbb{N}}$. For $\omega = \omega(0)\omega(1)\cdots\omega(n-1) \in \mathcal{A}^*$, we denote its length by $|\omega| = n$. By convention we set $|\varepsilon| = 0$.

A finite word w is called a *factor* of a finite (or an infinite) word u, denoted by $w \prec u$, if there exist a finite word p and a finite (or an infinite) word s such that u = pws. We say that the word w is a *prefix* of u, denoted by $w \triangleleft u$, if $p = \varepsilon$, and a *suffix* of u, denoted by $w \triangleright u$, if $s = \varepsilon$. Let $\mathbf{w} = w(0)w(1)w(2) \cdots \in \mathcal{A}^{\mathbb{N}}$ be an infinite word. Recall that for every positive integer n, the set of all factors of \mathbf{w} of length n is defined as

$$\mathcal{F}_{\mathbf{w}}(n) := \{ w(i)w(i+1)\cdots w(i+n-1) \mid i \ge 0 \}.$$

For convenience, we set $\mathcal{F}_{\mathbf{w}}(0) = \{\varepsilon\}$. Given a finite word $u = u(0)u(1)\cdots u(n-1)$ with $n \ge 2$, we denote the 1-length boundary word consisting of the first and last letter of u by ∂u , i.e., $\partial u = u(0)u(n-1)$. Define $\partial \mathcal{F}_{\mathbf{w}}(n) := \{\partial u \mid u \in \mathcal{F}_{\mathbf{w}}(n)\}$. For an integer $k \ge 2$ and a finite word $u = u(0)u(1)\cdots u(n) \in \mathcal{A}^*$ with $\mathcal{A} \subset \mathbb{N}$, let

$$u \equiv u'(0)u'(1)\cdots u'(n-1) \pmod{k}$$

where $u'(i) \in \Sigma_k$ and $u'(i) \equiv u(i) \pmod{k}$ for all *i*.

2.2. Binomial coefficients, k-binomial equivalence and k-binomial complexity

Now we introduce the binomial coefficients of words, the binomial equivalence of words and the binomial complexity of infinite words. Moreover, we list some properties of binomial coefficients of words. For more details, one can refer to [25].

Definition 1. (Binomial coefficient) Let \mathcal{A} be a non-empty finite set and $u = u(0)u(1)\cdots u(n-1) \in \mathcal{A}^n$. Let $s : \mathbb{N} \to \mathbb{N}$ be an increasing map such that $s(\ell-1) < n$ for all $1 \leq \ell \leq n$. Then for $\ell = 1, 2, \ldots, n$, the word $u(s(0)) \cdots u(s(\ell-1))$ is a scattered subword of length ℓ of u. The binomial coefficient $\binom{u}{v}$ of two finite words u and v is defined to be the number of times that v occurs as a scattered subword of u. In detail,

$$\binom{u}{v} = \sharp\{(i_1, i_2, \dots, i_{|v|}) \mid 0 \le i_1 < i_2 < \dots < i_{|v|} \le n - 1, u(i_1)u(i_2) \cdots u(i_{|v|}) = v\}.$$

For convenience, let $\binom{u}{v} = 0$ if |v| > |u| and $\binom{u}{\varepsilon} = 1$. For example, let $u = u(0)u(1)\cdots u(5) = 101000$ and v = 110. Then u(0)u(2)u(3) = u(0)u(2)u(4) = u(0)u(2)u(5) = v. Hence $\binom{u}{v} = 3$.

Definition 2. (k-binomial equivalence) Let k be a positive integer and let $\mathcal{A}^{\leq k}$ denote the set of words of length at most k over the alphabet \mathcal{A} . We say that $u, v \in \mathcal{A}^*$ are k-binomially equivalent if for every $x \in \mathcal{A}^{\leq k}$,

$$\binom{u}{x} = \binom{v}{x}.$$

We then write $u \sim_k v$ if u and v are k-binomially equivalent.

Indeed, since $\binom{u}{a} = |u|_a$ for all $a \in \mathcal{A}$, it is clear that $u \sim_1 v$ if and only if u and v are abelian equivalent. Note that, for all $k \geq 2$, if $u \sim_k v$, then $u \sim_\ell v$ for every $1 \leq \ell < k$.

There is an equivalent definition of the k-binominal equivalence using the extended Parikh vector. Let $u \in \Sigma_m^*$ and $1 \leq \ell \leq k$, there are exact m^{ℓ} words of length ℓ which can be enumerated lexicographically: $v_{\ell,1}, \ldots, v_{\ell,m^{\ell}}$. The extended Parikh vector of u, denoted by $\Psi_k(u)$, is

$$\Psi_k(u) := \left(\begin{pmatrix} u \\ v_{1,1} \end{pmatrix}, \dots, \begin{pmatrix} u \\ v_{1,m} \end{pmatrix}, \begin{pmatrix} u \\ v_{2,1} \end{pmatrix}, \dots, \begin{pmatrix} u \\ v_{2,m^2} \end{pmatrix}, \dots, \begin{pmatrix} u \\ v_{k,1} \end{pmatrix}, \dots, \begin{pmatrix} u \\ v_{k,m^k} \end{pmatrix} \right).$$

Then $u \sim_k v$ if and only if $\Psi_k(u) = \Psi_k(v)$. When k = 1, $\Psi_1(u)$ coincides with the Parikh vector of u. For convenience, we write $\Psi(u) = \Psi_1(u)$.

Example 1. Let u = 010001 and v = 001010. Then

$$\binom{u}{0} = 4, \binom{u}{1} = 2, \binom{u}{00} = 6, \binom{u}{01} = 5, \binom{u}{10} = 3, \binom{u}{11} = 1.$$

The same equalities hold for the word v. Therefore, $\Psi_2(u) = \Psi_2(v) = (4, 2, 6, 5, 3, 1)$ and $u \sim_2 v$.

Definition 3. (k-binomial complexity) Let \mathbf{w} be an infinite word and let k be a positive integer. The k-binomial complexity function $\mathbf{b}_{\mathbf{w},k}$ of \mathbf{w} is

$$b_{\mathbf{w},k}(n) := \sharp(\mathcal{F}_{\mathbf{w}}(n)/\sim_k) = \sharp\{\Psi_k(u) \mid u \in \mathcal{F}_{\mathbf{w}}(n)\}.$$

In the following, we collect some facts about binomial coefficients and binomial equivalence of words.

Lemma 1 ([25]). (1) Let u, v be two words and let a, a' be two letters. Then

$$\begin{pmatrix} au\\a'v \end{pmatrix} = \begin{pmatrix} u\\a'v \end{pmatrix} + \delta_{a,a'} \begin{pmatrix} u\\v \end{pmatrix} and \begin{pmatrix} ua\\va' \end{pmatrix} = \begin{pmatrix} u\\va' \end{pmatrix} + \delta_{a,a'} \begin{pmatrix} u\\v \end{pmatrix},$$

where $\delta_{a,a'} = 1$ if a = a' and 0 otherwise.

(2) Let s, w, t be three finite words over \mathcal{A} satisfying $|t| \leq |sw|$. Then

$$\binom{sw}{t} = \sum_{uv=t \text{ with } u, v \in \mathcal{A}^*} \binom{s}{u} \binom{w}{v}.$$

(3) Let u, v, w be three words and let $k \ge 1$ be an integer. Then

$$vu \sim_k wu \iff v \sim_k w \iff uv \sim_k uw$$

(4) Let $k \ge 2$ be an integer and let u, v, u', v' be four words such that $u \sim_{k-1} u'$ but $u \nsim_k u'$ and $v \sim_k v'$, then $uv \nsim_k u'v'$.

3. 2-binomial equivalence of factors of the generalized Thue-Morse word

Fix an integer $m \geq 3$. Recall that σ_m is the morphism over the alphabet $\Sigma_m = \{0, 1, \dots, m-1\}$ defined as $0 \mapsto 01 \cdots (m-1), 1 \mapsto 12 \cdots (m-1)0, \cdots, m-1 \mapsto (m-1)0 \cdots (m-2)$. The infinite word $\mathbf{t}_m := \sigma_m^{\infty}(0)$ is called the generalized Thue-Morse word. In this section, we investigate the 2-binomial equivalence between factors of the generalized Thue-Morse word.



Figure 1: Illustration for the intervals [2, 5] and (6, 1).

3.1. Properties of factors of the word \mathbf{t}_m .

We collect some basic properties for factors of \mathbf{t}_m .

Lemma 2. Let $\alpha \prec \sigma_m(a)$ and $\beta \prec \sigma_m(b)$ for some $a, b \in \Sigma_m$ with $\max\{|\alpha|, |\beta|\} < m$. Then $\alpha \sim_1 \beta$ if and only if $\alpha = \beta$.

Proof. Suppose $\alpha \sim_1 \beta$. By the definition of σ_m , we can assume that $\alpha = i(i+1)\cdots(i+|\alpha|-1)$ (mod m) and $\beta = j(j+1)\cdots(j+|\alpha|-1)$ (mod m) for some $i, j \in [0, m-1]$. If $i \neq j$, then there exists $x \in \Sigma_m$ such that $|\alpha|_x \neq |\beta|_x$, which means $\alpha \nsim_1 \beta$. Hence, i = j and $\alpha = \beta$.

For every $c, d \in \Sigma_m$ with $c \neq d$, write

$$[c,d] := \begin{cases} \{c,c+1,\ldots,d\}, & \text{if } c < d; \\ \{c,c+1,\ldots,m-1\} \cup \{0,1,\ldots,d\}, & \text{if } c > d. \end{cases}$$

Further, we write $(c, d) := [c, d] \setminus \{c, d\}$. Similarly, we define $(c, d] := [c, d] \setminus \{c\}$ and $[c, d] := [c, d] \setminus \{d\}$. For example, the intervals [2, 5] and (6, 1) are illustrated in Figure 1.

Lemma 3. Let $\gamma \prec \sigma_m(a)$ for some $a \in \Sigma_m$. Then for every $c, d \in \Sigma_m$ with $c \neq d$, we have

- (1) $|\gamma|_c, |\gamma|_d \in \{0, 1\}$. Moreover, if $|\gamma|_c = |\gamma|_d = 1$, then either $\binom{\gamma}{cd} = 1$ or $\binom{\gamma}{dc} = 1$.
- (2) If $\binom{\gamma}{cd} = 1$, then for every $e \in [c, d]$, we have $|\gamma|_e = 1$.
- (3) If $\binom{\gamma}{cd} = 1$ and $|\gamma| < m$, then $(d, c) \neq \emptyset$ and there exists $e \in (d, c)$ such that $|\gamma|_e = 0$.

Proof. The result follows directly from the definition of the morphism σ_m .

Lemma 4. Let $\alpha \prec \sigma_m(a)$, $\beta \prec \sigma_m(b)$, $\alpha' \prec \sigma_m(a')$ and $\beta' \prec \sigma_m(b')$ for some $a, b, a', b' \in \Sigma_m$. If $\alpha\beta \sim_1 \alpha'\beta'$ and there exist $c, d \in \Sigma_m$ with $c \neq d$ such that $\binom{\alpha}{cd} = \binom{\alpha'}{dc} = \binom{\beta'}{dc} = 1$, then $|\alpha| = m$.

Proof. Since $\binom{\alpha'}{dc} = \binom{\beta'}{dc} = 1$, by Lemma 3, for every $e \in [d, c]$, $|\alpha'|_e = |\beta'|_e = 1$. It follows from $\alpha\beta \sim_1 \alpha'\beta'$ that $|\alpha|_e = 1$ for all $e \in [d, c]$. Moreover, $\binom{\alpha}{cd} = 1$ implies that $|\alpha|_e = 1$ with $e \in [c, d]$. Hence, $|\alpha|_e = 1$ for every $e \in \Sigma_m$, which means $|\alpha| = m$.

3.2. 2-binomial equivalence between factors of \mathbf{t}_m

Let $c, d \in \Sigma_m$ with $c \neq d$ and $u \in \Sigma_m^*$. We have

$$\begin{pmatrix} \sigma_m(u) \\ cd \end{pmatrix} = \sum_{\substack{0 \le i < j \le |u| - 1}} |\sigma_m(u(i))|_c \cdot |\sigma_m(u(j))|_d + \sum_{\substack{0 \le i \le |u| - 1}} \binom{\sigma_m(u(i))}{cd}$$
$$= \binom{|u|}{2} + \sum_{x \in \Sigma_m, |u|_x \ge 1} |u|_x \binom{\sigma_m(x)}{cd}.$$

Note that for every $x \in \Sigma_m$, by the definition of σ_m ,

$$\begin{pmatrix} \sigma_m(x) \\ cd \end{pmatrix} = \begin{cases} 1, & x \in (d, c]; \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\binom{\sigma_m(u)}{cd} = \binom{|u|}{2} + \sum_{x \in (d,c]} |u|_x.$$
 (1)

Moreover, for all $\alpha, \beta \in \Sigma_m^*$,

$$\begin{pmatrix} \alpha \sigma_m(u)\beta\\cd \end{pmatrix} = \begin{pmatrix} \alpha\beta\\cd \end{pmatrix} + \begin{pmatrix} \sigma_m(u)\\cd \end{pmatrix} + |\alpha|_c |\sigma_m(u)|_d + |\sigma_m(u)|_c |\beta|_d = \begin{pmatrix} \alpha\beta\\cd \end{pmatrix} + |u|(|\alpha|_c + |\beta|_d) + \sum_{x \in (d,c]} |u|_x + \begin{pmatrix} |u|\\2 \end{pmatrix}$$
(2)

where $0 \leq \sum_{x \in (d,c]} |u|_x \leq |u|$. Recall that the sets S and \mathcal{P} are defined as

$$\mathcal{S} := \{ w \in \Sigma_m^* \mid w \triangleright \sigma_m(a) \text{ for some } a \in \Sigma_m \text{ and } |w| < m \},\$$
$$\mathcal{P} := \{ w \in \Sigma_m^* \mid w \triangleleft \sigma_m(a) \text{ for some } a \in \Sigma_m \text{ and } |w| < m \}.$$

Now we give the proof of Theorem 2.

Proof of Theorem 2. Since $\max\{|\alpha| + |\beta|, |\alpha'| + |\beta'|\} \le 2m - 2 < 2m$, we have $||u'| - |v'|| \le 1$. Without loss of generality, we may assume $|u'| \ge |v'|$. The proof of Theorem 2 is separated into the following cases according to the lengths of α , β , α' , β' , u' and v':

- 1. |u'| = |v'| and $|\alpha| + |\beta| = 0$, see Proposition 1;
- 2. |u'| = |v'| and $|\alpha| + |\beta| \ge 1$, see Propositions 2 and 3;
- 3. |u'| = |v'| + 1 and $|\alpha| + |\beta| = 0$, see Proposition 4;
- 4. |u'| = |v'| + 1, $|\alpha| = 0$ and $|\beta| \ge 1$, see Proposition 5;
- 5. |u'| = |v'| + 1, $|\beta| = 0$ and $|\alpha| \ge 1$, see Proposition 6;
- 6. |u'| = |v'| + 1, $|\alpha| \ge 1$ and $|\beta| \ge 1$, see Proposition 7.

Proposition 1. For every $u', v' \in \Sigma_m^*$, $\sigma_m(u') \sim_2 \sigma_m(v')$ if and only if $u' \sim_1 v'$.

Proof. If $u' \sim_1 v'$, then $|\sigma_m(u')|_a = |u'| = |v'| = |\sigma_m(v')|_a$ for all $a \in \Sigma_m$. So, for all $a \in \Sigma_m$,

$$\binom{\sigma_m(u')}{aa} = \binom{|\sigma_m(u')|_a}{2} = \binom{|u'|}{2} = \binom{|v'|}{2} = \binom{|\sigma_m(v')|_a}{2} = \binom{\sigma_m(v')}{aa}.$$

Moreover, for any $c, d \in \Sigma_m$ with $c \neq d$, since $u' \sim_1 v'$, by (1),

$$\binom{\sigma_m(u')}{cd} = \binom{|u'|}{2} + \sum_{x \in (d,c]} |u'|_x = \binom{|v'|}{2} + \sum_{x \in (d,c]} |v'|_x = \binom{\sigma_m(v')}{cd}.$$

Therefore, $\sigma_m(u') \sim_2 \sigma_m(v')$.

Conversely, if $\sigma_m(u') \sim_2 \sigma_m(v')$, then |u'| = |v'| and $\binom{\sigma(u')}{a(a-1)} = \binom{\sigma(v')}{a(a-1)}$ for all $a \in \Sigma_m \setminus \{0\}$. Note that $((a-1), a] = \{a\}$. It follows from (1) that

$$\binom{\sigma_m(u')}{a(a-1)} = \binom{|u'|}{2} + \sum_{x \in \{a\}} |u'|_x = \binom{|u'|}{2} + |u'|_a$$

and

$$\binom{\sigma_m(v')}{a(a-1)} = \binom{|v'|}{2} + \sum_{x \in \{a\}} |v'|_x = \binom{|v'|}{2} + |v'|_a.$$

Thus, $|u'|_a = |v'|_a$ for every $a \in \Sigma_m \setminus \{0\}$. Since |u'| = |v'|, we have $u' \sim_1 v'$.

Proposition 2. Let (α, β) , $(\alpha', \beta') \in S \times P$ with $|\alpha| + |\beta| \ge 1$. Let $u, v \in \mathcal{F}_{\mathbf{t}_m}(n)$ with $n \ge 3$. If $\alpha \sigma_m(u)\beta \sim_2 \alpha' \sigma_m(v)\beta'$ and $\alpha \beta \sim_2 \alpha' \beta'$, then $\alpha = \alpha', \beta = \beta'$ and $u \sim_1 v$.

Proof. Since $\alpha\beta \sim_2 \alpha'\beta'$, |u| = |v| and $\alpha\sigma_m(u)\beta \sim_2 \alpha'\sigma_m(v)\beta'$, it follows from (2) that for every $c \in \Sigma_m \setminus \{0\}$,

$$0 = \begin{pmatrix} \alpha \sigma_m(u)\beta\\c(c-1) \end{pmatrix} - \begin{pmatrix} \alpha \sigma_m(v)\beta\\c(c-1) \end{pmatrix}$$
$$= |u| (|\alpha|_c + |\beta|_{c-1} - |\alpha'|_c - |\beta'|_{c-1}) + |u|_c - |v|_c$$

This implies that $|v|_c - |u|_c$ is divisible by |u|. Note that $|u| = |v| \ge 3$ and the word \mathbf{t}_m is cube-free. This implies $-|u| < |v|_c - |u|_c < |u|$. Therefore, $|u|_c = |v|_c$ for all $c \in \Sigma_m \setminus \{0\}$. Note also that |u| = |v|, we have $u \sim_1 v$.

For any $c, d \in \Sigma_m$, by (2), we have

$$\begin{pmatrix} \alpha \sigma_m(u)\beta \\ cd \end{pmatrix} = \begin{pmatrix} \alpha\beta \\ cd \end{pmatrix} + |u|(|\alpha|_c + |\beta|_d) + \begin{pmatrix} \sigma_m(u) \\ cd \end{pmatrix}$$
(3)

and

$$\begin{pmatrix} \alpha' \sigma_m(v)\beta'\\ cd \end{pmatrix} = \begin{pmatrix} \alpha'\beta'\\ cd \end{pmatrix} + |v|(|\alpha'|_c + |\beta'|_d) + \begin{pmatrix} \sigma_m(v)\\ cd \end{pmatrix}.$$
(4)

Since $u \sim_1 v$, by Proposition 1, we have $\sigma_m(u) \sim_2 \sigma_m(v)$ and $\binom{\sigma_m(u)}{cd} = \binom{\sigma_m(v)}{cd}$. Further, since $\alpha \sigma_m(u) \beta \sim_2 \alpha' \sigma_m(v) \beta'$ and $\alpha \beta \sim_2 \alpha' \beta'$, we have

$$\begin{pmatrix} \alpha' \sigma_m(u)\beta' \\ cd \end{pmatrix} = \begin{pmatrix} \alpha \sigma_m(v)\beta \\ cd \end{pmatrix} \text{ and } \begin{pmatrix} \alpha\beta \\ cd \end{pmatrix} = \begin{pmatrix} \alpha'\beta' \\ cd \end{pmatrix}$$

Note that |u| = |v|. It follows from (3) and (4) that for any $c, d \in \Sigma_m$,

$$|\alpha|_c + |\beta|_d = |\alpha'|_c + |\beta'|_d.$$

$$\tag{5}$$

Now we prove $\alpha = \alpha'$ and $\beta = \beta'$. There are two sub-cases.

(i) $|\alpha| \ge 1$ and $|\beta| \ge 1$. For any $c, d \in \Sigma_m$ satisfying $|\alpha|_c = |\beta|_d = 1$, by (5), we have

$$|\alpha'|_c + |\beta'|_d = 2$$

Since $(\alpha', \beta') \in \mathcal{S} \times \mathcal{P}$, we see that $|\alpha'|_c = |\beta'|_d = 1$. In the same way, one can show that $|\alpha|_c = |\beta|_d = 1$ for any $c, d \in \Sigma_m$ satisfying $|\alpha'|_c = |\beta'|_d = 1$. This implies that $\alpha \sim_1 \alpha'$ and $\beta \sim_1 \beta'$. By Lemma 2, $\alpha = \alpha'$ and $\beta = \beta'$.

(ii) $|\alpha| = 0$ or $|\beta| = 0$. It suffices to show the case $|\alpha| = 0$. If $\beta = d$ for some $d \in \Sigma_m$, then $\alpha'\beta' \sim_2 \alpha\beta$ implies that $\alpha'\beta' = d$. For any $c \in \Sigma_m$ with $c \neq d$, by (5), $1 = |\alpha'|_c + |\beta'|_d$. So, $\alpha' = \varepsilon$ and $\beta' = d$. If $|\beta| > 1$, then there exists a $d \in \Sigma_m$ with $|\beta|_d = 1$. For any $c \in \Sigma_m$ with $c \neq d$, by (5), we have

$$|\alpha'|_{c} + |\beta'|_{d} = |\beta|_{d} = 1.$$
(6)

Since $\alpha\beta \sim_2 \alpha'\beta'$, which means $\alpha\beta \sim_1 \alpha'\beta'$, we have

$$|\alpha'|_d + |\beta'|_d = |\beta|_d = 1.$$
 (7)

Combining (6) and (7), we obtain that for all $x \in \Sigma_m$, $|\alpha'|_x = 1 - |\beta'|_d$. If $|\beta'|_d = 0$, then $|\alpha'| = m$ which contradicts to the fact $|\alpha'| < m$. If $|\beta'|_d = 1$, then $\alpha' = \varepsilon$. Therefore, $\beta = \alpha\beta \sim_1 \alpha'\beta' = \beta'$.

Proposition 3. Let (α, β) , $(\alpha', \beta') \in \mathcal{S} \times \mathcal{P}$ with $|\alpha| + |\beta| \ge 1$. If $\alpha\beta \nsim_2 \alpha'\beta'$, then for any $u, u' \in \Sigma_m^*$ with |u| = |u'|, $\alpha\sigma_m(u)\beta \nsim_2 \alpha'\sigma_m(u')\beta'$.

Proof. Since |u| = |u'|, we have $\sigma_m(u) \sim_1 \sigma_m(u')$. If $\alpha\beta \nsim_1 \alpha'\beta'$, then $\alpha\sigma_m(u)\beta \nsim_1 \alpha'\sigma_m(u')\beta'$. In the rest, we assume that $\alpha\beta \sim_1 \alpha'\beta'$.

Since $\alpha\beta \approx_2 \alpha'\beta'$ and $\alpha\beta \sim_1 \alpha'\beta'$, there exist $c, d \in \Sigma_m$ with $c \neq d$ such that $\binom{\alpha\beta}{cd} \neq \binom{\alpha'\beta'}{cd}$. Note that

$$\begin{pmatrix} \alpha\beta\\ cd \end{pmatrix} = \begin{pmatrix} \alpha\\ cd \end{pmatrix} + \begin{pmatrix} \beta\\ cd \end{pmatrix} + |\alpha|_c |\beta|_d.$$
(8)

This implies that $\binom{\alpha\beta}{cd}, \binom{\alpha'\beta'}{cd} \in \{0, 1, 2, 3\}$. Without loss of generality, we can assume that $\binom{\alpha\beta}{cd} > \binom{\alpha'\beta'}{cd}$. Then, $\binom{\alpha\beta}{cd} \in \{1, 2, 3\}$. If $\binom{\alpha\beta}{cd} = 3$, then it follows from (8) that $\binom{\alpha}{cd} = \binom{\beta}{cd} = 1$ and $|\alpha\beta|_c = |\alpha\beta_d| = 2$. Since $\alpha\beta \sim_1 \alpha'\beta'$, we have $|\alpha'\beta'|_c = |\alpha'\beta'|_d = 2$. Note also that $\binom{\alpha'\beta'}{cd} < 3$, by Lemma 3, at least one of $\binom{\alpha'}{dc}$ and $\binom{\beta'}{dc}$ equals 1. By Lemma 4, $|\alpha'| = m$ or $|\beta'| = m$ which contradicts to the fact that $(\alpha', \beta') \in \mathcal{S} \times \mathcal{P}$. So $\binom{\alpha\beta}{cd} \neq 3$.

In the following, we deal with the cases that $\binom{\alpha\beta}{cd} = 1$ or 2. By (2),

$$\begin{pmatrix} \alpha \sigma_m(u)\beta\\cd \end{pmatrix} \ge \begin{pmatrix} \alpha\beta\\cd \end{pmatrix} + |u|(|\alpha|_c + |\beta|_d) + \begin{pmatrix} |u|\\2 \end{pmatrix}, \\ \begin{pmatrix} \alpha' \sigma_m(u')\beta'\\cd \end{pmatrix} \le \begin{pmatrix} \alpha'\beta'\\cd \end{pmatrix} + |u'|(|\alpha'|_c + |\beta'|_d) + \begin{pmatrix} |u'|\\2 \end{pmatrix} + |u'|.$$

Since |u| = |u'| and $\binom{\alpha\beta}{cd} > \binom{\alpha'\beta'}{cd}$, we have

$$\binom{\alpha \sigma_m(u)\beta}{cd} - \binom{\alpha' \sigma_m(u')\beta'}{cd} > |u| (|\alpha|_c + |\beta|_d - |\alpha'|_c - |\beta'|_d - 1).$$
(9)

When $\binom{\alpha\beta}{cd} = 2$, by (8), we have $|\alpha|_c + |\beta|_d = 2$. If $|\alpha'|_c |\beta'|_d = 0$, then $|\alpha'|_c + |\beta'|_d \leq 1$. It follows from (9) that $\binom{\alpha\sigma_m(u)\beta}{cd} > \binom{\alpha'\sigma_m(u')\beta'}{cd}$ and $\alpha\sigma_m(u)\beta \approx_2 \alpha'\sigma_m(u')\beta'$. If $|\alpha'|_c |\beta'|_d = 1$, then $2 = \binom{\alpha\beta}{cd} > \binom{\alpha'\beta'}{cd}$ implies $\binom{\alpha'}{cd} = \binom{\beta'}{cd} = 0$. Since $\binom{\alpha\beta}{cd} = 2$, following (8), we see that either $\binom{\alpha}{cd} = 1$ or $\binom{\beta}{cd} = 1$.

- 1. If $\binom{\alpha}{cd} = 1$, then $|\alpha\beta|_d = 2$. Since $\alpha\beta \sim_1 \alpha'\beta'$, we have $|\alpha'\beta'|_d = 2$ which implies $|\alpha'|_d = 1$. Recall that $|\alpha'|_c = 1$ and $\binom{\alpha'}{cd} = 0$. We have $\binom{\alpha'}{dc} = 1$. By Lemma 3, there exists an $e \in (c, d) \neq \emptyset$ such that $|\alpha'|_e = 0$ and $|\alpha|_e = 1$. Moreover, $\alpha\beta \sim_1 \alpha'\beta'$ yields that $|\beta'|_e = 1$ and $|\beta|_e = 0$. Then $\binom{\alpha\beta}{ed} = 2 > \binom{\alpha'\beta'}{ed}$ and $|\alpha|_e + |\beta|_d - |\alpha'|_e - |\beta'|_d = 1 + 1 - 0 - 1 = 1$. Applying (9), we have $\binom{\alpha\sigma_m(u)\beta}{ed} > \binom{\alpha'\sigma_m(u')\beta'}{ed}$.
- 2. If $\binom{\beta}{cd} = 1$, then $|\alpha\beta|_c = 2$. It follows from $\alpha\beta \sim_1 \alpha'\beta'$ and $\binom{\beta'}{cd} = 0$ that $\binom{\beta'}{dc} = 1$. By Lemma 3, there exists an $e \in (c,d) \neq \emptyset$ such that $|\beta'|_e = 0$ and $|\beta|_e = 1$. Moreover, $\alpha\beta \sim_1 \alpha'\beta'$ yields that $|\alpha'|_e = 1$ and $|\alpha|_e = 0$. Then $\binom{\alpha\beta}{ce} = 2 > \binom{\alpha'\beta'}{ce'}$ and $|\alpha|_c + |\beta|_e |\alpha'|_c |\beta'|_e = 1 + 1 1 0 = 1$. Applying (9), we have $\binom{\alpha\sigma_m(u)\beta}{ce} > \binom{\alpha'\sigma_m(u')\beta'}{ce}$.

When $\binom{\alpha\beta}{cd} = 1$, we have $\binom{\alpha'\beta'}{cd} = 0$. Then $|\alpha|_c + |\beta|_d \ge 1 \ge |\alpha'|_c + |\beta'|_d$. If $(|\alpha|_c + |\beta|_d) - (|\alpha'|_c + |\beta'|_d) \ge 1$, then by (9), we have $\binom{\alpha\sigma_m(u)\beta}{cd} > \binom{\alpha'\sigma_m(u')\beta'}{cd}$. Now suppose that $|\alpha|_c + |\beta|_d = |\alpha'|_c + |\beta'|_d = 1$. There are four sub-cases.

1. $|\alpha|_c = 1, |\beta|_d = 0$. Then, it follows from $\binom{\alpha\beta}{cd} = 1$ that $|\alpha|_d = 1$ and $\binom{\alpha}{cd} = 1$.

- If $|\alpha'|_c = 1$ and $|\beta'|_d = 0$, then $\alpha\beta \sim_1 \alpha'\beta'$ implies that $|\alpha'|_d = 1$ and $\binom{\alpha'}{dc} = 1$. By Lemma 3, there exists an $e \in (c,d) \neq \emptyset$ such that $|\alpha'|_e = 0$ and $|\alpha|_e = 1$. Then $\binom{\alpha\beta}{ed} = 1 > \binom{\alpha'\beta'}{ed} = 0$ and $|\alpha|_e + |\beta|_d |\alpha'|_e |\beta'|_d = 1 + 0 0 0 = 1$. Applying (9), we have $\binom{\alpha\sigma_m(u)\beta}{ed} > \binom{\alpha'\sigma_m(u')\beta'}{ed}$.
- If $|\alpha'|_c = 0$ and $|\beta'|_d = 1$. Since $\alpha\beta \sim_1 \alpha'\beta'$, we have $|\beta'|_c = 1$ and $\binom{\beta'}{dc} = 1$. By Lemma 3, there exists an $e \in (c,d) \neq \emptyset$ such that $|\beta'|_e = 0$ and $|\alpha|_e = 1$. Moreover, $\alpha\beta \sim_1 \alpha'\beta'$ yields that $|\alpha'|_e = 1$ and $|\beta|_e = 0$. Then $\binom{\alpha\beta}{ce} = 1 > \binom{\alpha'\beta'}{ce} = 0$ and $|\alpha|_c + |\beta|_e |\alpha'|_c |\beta'|_e = 1 + 0 0 0 = 1$. Applying (9), we have $\binom{\alpha\sigma_m(u)\beta}{ce} > \binom{\alpha'\sigma_m(u')\beta'}{ce}$.
- 2. $|\alpha|_c = 0, |\beta|_d = 1$. It follows from $\binom{\alpha\beta}{cd} = 1$ that $|\beta|_c = 1$ and $\binom{\beta}{cd} = 1$.
 - If $|\alpha'|_c = 1$ and $|\beta'|_d = 0$, then $\alpha\beta \sim_1 \alpha'\beta'$ yields $|\alpha|_d = 0$, $|\alpha'|_d = 1$ and $\binom{\alpha'}{dc} = 1$. By Lemma 3, there exists an $e \in (d, c)$ such that $|\alpha'|_e = 1$ and $|\beta|_e = 0$. Since $\alpha\beta \sim_1 \alpha'\beta'$, we have $|\beta'|_e = 0$ and $|\alpha|_e = 1$. Then $\binom{\alpha'\beta'}{de} = 1 > \binom{\alpha\beta}{de} = 0$ and $|\alpha'|_d + |\beta'|_e - |\alpha|_d - |\beta|_e = 1 + 0 - 0 - 0 = 1$. Applying (9), we have $\binom{\alpha'\sigma_m(u')\beta'}{de} > \binom{\alpha\sigma_m(u)\beta}{de}$.
 - If $|\alpha'|_c = 0$ and $|\beta'|_d = 1$, then $\binom{\beta'}{dc} = 1$. Using a similar argument, there exists an $e \in (c,d)$ such that $|\beta|_e = 1$ and $|\beta'|_e = 0$. Consequently, $\binom{\alpha \sigma_m(u)\beta}{ce} > \binom{\alpha' \sigma_m(u')\beta'}{ce}$.

Proposition 4. Let $(\alpha, \beta) \in S \times P$ and $\alpha, \beta \neq \varepsilon$. Then for any $u \in \mathcal{F}_{\mathbf{t}_m}(n)$ and $v \in \mathcal{F}_{\mathbf{t}_m}(n+1)$ with $n \geq 3$, $\alpha \sigma_m(u)\beta \nsim_2 \sigma_m(v)$.

Proof. Suppose that v = v'c where $c \in \Sigma_m$. Since |u| = |v'|, we have $\sigma_m(u) \sim_1 \sigma_m(v')$. If $\alpha\beta \approx_1 \sigma_m(c)$, then $\alpha\sigma_m(u)\beta \approx_1 \sigma_m(v)$. Now we assume that $\alpha\beta \sim_1 \sigma_m(c)$. There are two cases.

(i) $\alpha\beta = \sigma_m(c)$. Let d be the last letter of α and let e be the first letter of β . By the definition of σ_m , $e \equiv d+1 \pmod{m}$ and $|\alpha|_e = |\beta|_d = 0$. By (2),

$$\binom{\alpha \sigma_m(u)\beta}{ed} = \binom{\sigma_m(u)}{ed} = \binom{|u|}{2} + \sum_{x \in (d,e]} |u|_x = \binom{|u|}{2} + |u|_e$$

and

$$\binom{\sigma_m(v'c)}{ed} = \binom{\sigma_m(v')}{ed} + |v'| = \binom{|v'|}{2} + |v'|_e + |v'|_e$$

Since the word \mathbf{t}_m is cube-free and $|u| \ge 3$, $|u|_e < |u| = |v'|$. Hence $\binom{\alpha \sigma_m(u)\beta}{ed} < \binom{\sigma_m(v'c)}{ed}$ and $\alpha \sigma_m(u)\beta \nsim_2 \sigma_m(v'c)$.

(ii) $\alpha\beta \neq \sigma_m(c)$. There exist $d, e \in \Sigma_m$ such that $|\alpha|_d = 1 = |\beta|_e$ and $\binom{\sigma_m(c)}{de} = 0$. Note that $|\alpha|_e = |\beta|_d = 0$. By (2),

$$\binom{\alpha \sigma_m(u)\beta}{de} \ge 1 + 2|u| + \binom{|u|}{2} > 2|v'| + \binom{|v'|}{2} \ge \binom{\sigma_m(v'c)}{de}.$$

Therefore $\alpha \sigma_m(u) \beta \not\sim_2 \sigma_m(v'c)$.

Proposition 5. Let β , $\beta' \in \mathcal{P}$ and $\alpha' \in \mathcal{S}$ where β , α' and β' are nonempty words. Then for any $u \in \mathcal{F}_{\mathbf{t}_m}(n+1)$ and $v \in \mathcal{F}_{\mathbf{t}_m}(n)$ with $n \geq 3$, we have $\sigma_m(u)\beta \nsim_2 \alpha' \sigma_m(v)\beta'$.

Proof. Let u = au' where $a \in \Sigma_m$. If $\alpha'\beta' \approx_1 \sigma_m(a)\beta$, then $\sigma_m(au')\beta \approx_1 \alpha'\sigma_m(v)\beta'$. In the following, we assume that $\sigma_m(a)\beta \sim_1 \alpha'\beta'$. Then, for every $x \in \Sigma_m$ with $|\beta|_x = 1$, we have $|\alpha'|_x = |\beta'|_x = 1$. Noting that $\sigma_m(a)\beta \sim_1 \alpha'\beta'$, we see that $|\beta'| > |\beta|$.

Let b and c be the first and the last letter of β respectively. (If $|\beta| = 1$, then b = c.) Since $\sigma_m(a)\beta \sim_1 \alpha'\beta'$, we have $|\alpha'|_b = |\beta'|_b = |\alpha'|_c = |\beta'|_c = 1$.

- If b is the first letter of β' , then c is not the last letter of β' . Moreover, letting $d \equiv c+1$ (mod m), we have $|\beta|_d = 0$ and $|\beta'|_d = 1$. It follows from $\sigma_m(a)\beta \sim_1 \alpha'\beta'$ that $|\alpha'|_d = 0$. Then $\binom{\sigma_m(a)\beta}{cd} \leq 1 < 2 = \binom{\alpha'\beta'}{cd}$ and $|\alpha'|_c + |\beta'|_d - |\sigma_m(a)|_c - |\beta|_d = 1 + 1 - 1 - 0 = 1$. Using (9), we have $\binom{\sigma_m(au')\beta}{cd} < \binom{\alpha'\sigma_m(v)\beta'}{cd}$ and $\sigma_m(au')\beta \not\sim_2 \alpha'\sigma_m(v)\beta'$.
- If b is not the first letter of β' , then $e \equiv b-1 \pmod{m}$ occurs in β' and $|\alpha'|_e = 0$. Now $\binom{\sigma_m(a)\beta}{eb} \geq 1 = \binom{\alpha'\beta'}{eb}$ and $|\sigma_m(a)|_e + |\beta|_b |\alpha'|_e |\beta'|_b = 1 + 1 0 1 = 1$. By the fact that the word \mathbf{t}_m is cube-free and $|u'| \geq 3$, we have $\sum_{x \in (b,e]} |u'|_x = |u'| |u'|_b > 0$. Using (2), we have $\binom{\sigma_m(au')\beta}{eb} \geq \binom{\alpha'\sigma_m(v)\beta'}{eb}$ and $\sigma_m(au')\beta \not\approx_2 \alpha'\sigma_m(v)\beta'$.

Proposition 6. Let $\alpha, \alpha' \in S$ and $\beta' \in P$ where α , α' and β' are nonempty. Then for any $u \in \mathcal{F}_{\mathbf{t}_m}(n+1)$ and $v \in \mathcal{F}_{\mathbf{t}_m}(n)$ with $n \geq 3$, we have $\alpha \sigma_m(u) \nsim_2 \alpha' \sigma_m(v) \beta'$.

Proof. Let u = u'b for some $b \in \Sigma_m$. If $\alpha \sigma_m(b) \approx_1 \alpha' \beta'$, then $\alpha \sigma_m(u) \approx_1 \alpha' \sigma_m(v) \beta'$. Now we assume that $\alpha \sigma_m(b) \sim_1 \alpha' \beta'$. Then $|\alpha'| > |\alpha|$ and for every $x \in \Sigma_m$ with $|\alpha|_x = 1$, $|\alpha'|_x = |\beta'|_x = 1$. Let c and d be the first and last letter of α respectively. (If $|\alpha| = 1$, then c = d.)

• If c is the first letter of α' , then set $e \equiv d+1 \pmod{m}$. Consequently, $\binom{\alpha'}{de} = 1$ and $|\alpha|_e = 0$. By $\alpha \sigma_m(b) \sim_1 \alpha' \beta'$, we have $|\beta'|_e = 0$. Note that $\binom{\alpha \sigma_m(b)}{de} \geq 1 = \binom{\alpha' \beta'}{de}$, $|\alpha|_d + |\sigma_m(b)|_e - |\alpha'|_d - |\beta'|_e = 1 + 1 - 1 - 0 = 1$. It follows from the word \mathbf{t}_m is cube-free and $|u| \geq 3$ that $\sum_{x \in (e,d]} |u'|_x = |u'| - |u'|_e > 0$. By (2) and |u'| = |v|, we have

$$\binom{\alpha \sigma_m(u'b)}{de} > 1 + 2|u'| + \binom{|u'|}{2} \ge \binom{\alpha' \sigma_m(v)\beta'}{de}.$$

Hence $\alpha \sigma_m(u'b) \not\sim_2 \alpha' \sigma_m(v) \beta'$.

• If c is not the first letter of α' , then set $e \equiv c-1 \pmod{m}$. Then $\binom{\alpha'}{ec} = 1$ and $|\alpha|_e = 0$. Now $\binom{\alpha'\beta'}{ec} = 2 > 1 \ge \binom{\alpha\sigma_m(b)}{ec}$ and $|\alpha'|_e + |\beta'|_c - |\alpha|_e - |\sigma_m(b)|_c = 1 + 1 - 0 - 1 = 1$. By (9), $\binom{\alpha'\sigma_m(v)\beta'}{ec} > \binom{\alpha\sigma_m(u'b)}{ec}$. So, $\alpha\sigma_m(u'b) \not\approx_2 \alpha'\sigma_m(v)\beta'$.

Proposition 7. Let $(\alpha, \beta), (\alpha', \beta') \in S \times P$ and $\min\{|\alpha|, |\beta|, |\alpha'|, |\beta'|\} \geq 1$. Then for any $u \in \mathcal{F}_{\mathbf{t}_m}(n+1), v \in \mathcal{F}_{\mathbf{t}_m}(n) \text{ with } n \geq 3, \text{ we have } \alpha \sigma_m(u) \beta \nsim_2 \alpha' \sigma_m(v) \beta'.$

Proof. Let u = u'c for some $c \in \Sigma_m$. If $\alpha \sigma_m(c)\beta \not\sim_1 \alpha'\beta'$, then $\alpha \sigma_m(u'c)\beta \not\sim_1 \alpha'\sigma_m(v)\beta'$. From now on, we assume that $\alpha \sigma_m(c)\beta \sim_1 \alpha'\beta'$. Then for all $x \in \Sigma_m$, $|\alpha|_x + |\beta|_x = |\alpha \sigma_m(c)\beta|_x - 1 =$ $|\alpha'\beta'|_x - 1 \leq 1$. Consequently, there exist $d, e \in \Sigma_m$ with $d \neq e$ such that $|\alpha|_d = |\beta|_e = 1$. Hence $|\beta|_d = |\alpha|_e = 0$ and $|\alpha'|_d = |\beta'|_d = |\alpha'|_e = |\beta'|_e = 1$. By (2),

$$\binom{\alpha \sigma_m(u)\beta}{de} \ge \binom{\alpha \beta}{de} + |u|(|\alpha|_d + |\beta|_e) + \binom{|u|}{2} = 3 + 3|u'| + \binom{|u'|}{2}$$

and

$$\binom{\alpha'\sigma_m(v)\beta'}{de} \le \binom{\alpha'\beta'}{de} + 3|v| + \binom{|v|}{2} = \binom{\alpha'\beta'}{de} + 3|u'| + \binom{|u'|}{2}.$$

Note that $\binom{\alpha'\beta'}{de} \leq 3$. If $\binom{\alpha'\beta'}{de} < 3$, then $\binom{\alpha\sigma_m(u)\beta}{de} > \binom{\alpha'\sigma_m(v)\beta'}{de}$ and $\alpha\sigma_m(u)\beta \nsim_2 \alpha'\sigma_m(v)\beta'$. If $\binom{\alpha'\beta'}{de} = 3$, then $\binom{\alpha'}{de} = \binom{\beta'}{de} = 1$. For every $x \in [d, e]$, $|\alpha'|_x = |\beta'|_x = 1$ and $|\alpha\beta|_x = 1$ since $\alpha\sigma_m(c)\beta \sim_1 \alpha'\beta'$. Then, it follows from $(\alpha, \beta) \in \mathcal{S} \times \mathcal{P}$ and $|\alpha|_d = |\beta|_e = 1$ that there exists $y \in [d, e)$ such that $y \triangleright \alpha$ and $z \triangleleft \beta$ where $z \equiv (y+1) \pmod{m}$. As the word \mathbf{t}_m is cube-free and |u| > 3, we have $\sum_{x \in (z, y]} |u|_x = |u| - |u|_z > 0$. By (2),

$$\begin{pmatrix} \alpha \sigma_m(u)\beta \\ yz \end{pmatrix} > \begin{pmatrix} \alpha \beta \\ yz \end{pmatrix} + |u|(|\alpha|_y + |\beta|_z) + \begin{pmatrix} |u| \\ 2 \end{pmatrix} = 3 + 3|u'| + \begin{pmatrix} |u'| \\ 2 \end{pmatrix} \ge \begin{pmatrix} \alpha' \sigma_m(v)\beta' \\ yz \end{pmatrix}.$$

$$(e, \alpha \sigma_m(u)\beta \not\sim_2 \alpha' \sigma_m(v)\beta'.$$

Hence, $\alpha \sigma_m(u) \beta \not\sim_2 \alpha' \sigma_m(v) \beta'$.

4. 2-binomial complexity of the generalized Thue-Morse word

The aim of this section is to compute the 2-binomial complexity of the generalized Thue-Morse word \mathbf{t}_m . For every $u \in \Sigma_m^*$, the Parikh vector of u is denoted by

 $\Psi(u) := (|u|_0, |u|_1, \cdots, |u|_{m-1}).$

Write $\mathbf{1} := (1, 1, \dots, 1) = \Psi(\sigma_m(0))$. For $x \in \Sigma_m$ and $\ell \in \mathbb{N}$, define $x|_{\ell} \in \Sigma_m^{\ell}$ as follow

$$x|_{\ell} \equiv x(x+1)\dots(x+\ell-1) \pmod{m}.$$

If $\ell = 0$, then $x|_{\ell} := \varepsilon$.

To compute the 2-binomial complexity $b_{\mathbf{t}_m,2}(n)$, we need some auxiliary lemmas.

Lemma 5 (Lemma 2 in [4]). For the generalized Thue-Morse sequence \mathbf{t}_m and every integer $n \geq 2$, we have $\partial \mathcal{F}_{\mathbf{t}_m}(n) = \Sigma_m^2$.

Lemma 6. Let n = km with $k \ge 1$. Then for every $a \in \Sigma_m$, we have

$$\sharp\{\Psi(au)\mid au\in\mathcal{F}_{\mathbf{t}_m}(n)\}=\sharp\{\Psi(ua)\mid ua\in\mathcal{F}_{\mathbf{t}_m}(n)\}=1+\frac{m(m-1)}{2}.$$

Proof. For $i \in [0, m-1]$ and $b \in \Sigma_m$, write

$$\mathcal{E}_{i,b} := \{ \alpha \sigma_m(v) \beta \in \mathcal{F}_{\mathbf{t}_m}(n) \mid \alpha = a \mid_i, \beta = b \mid_{m-i}, v \in \mathcal{F}_{\mathbf{t}_m}(k-1), a \triangleleft \alpha \sigma_m(v) \}.$$

By Lemma 5, for every $i \in [0, m-1]$ and $b \in \Sigma_m$, there exists some $v \in \Sigma_m^*$ such that $\mathcal{E}_{i,b} \neq \emptyset$. Let $\mathcal{E}_i = \bigcup_{b \in \Sigma_m} \mathcal{E}_{i,b}$. Then $\{w \in \mathcal{F}_{\mathbf{t}_m}(n) \mid a \triangleleft w\} = \bigcup_{j=0}^{m-1} \mathcal{E}_j$. Note that

$$E_{i,b} := \{ \Psi(w) \mid w \in \mathcal{E}_{i,b} \} = \{ \Psi(a|_i) + \Psi(b|_{m-i}) + \Psi(\sigma_m(v)) \mid v \in \mathcal{F}_{\mathbf{t}_m}(k-1) \}$$

= $\{ \Psi(a|_i) + \Psi(b|_{m-i}) + (k-1)\mathbf{1} \}.$

Write $E_i := \{\Psi(w) \mid w \in \mathcal{E}_i\}$. Then $\sharp E_0 = 1$ and for $i \in [1, m - 1]$, $\sharp E_i = m$. Observe that for any $i \in [1, m - 1]$ and $b \in \Sigma_m$,

$$\begin{cases} E_{i,b} = E_0, & \text{if } b - a - i \equiv 0 \pmod{m}; \\ E_{i,b} = E_{j,b}, & \text{if } b - a - i \equiv j \in [1, i - 1]; \\ E_{i,b} \cap \left(\bigcup_{\ell=0}^{i-1} E_{\ell} \right) = \emptyset, & \text{otherwise.} \end{cases}$$

So, for $i \in [1, m-1], \sharp \left(E_i \setminus \left(\bigcup_{j=0}^{i-1} E_j\right)\right) = m-i$. Hence

$$\sharp\{\Psi(au) \mid au \in \mathcal{F}_{\mathbf{t}_m}(n)\} = \sharp\left(\bigcup_{j=0}^{m-1} E_i\right) = 1 + (m-1) + (m-2) + \dots + 1 = 1 + \frac{m(m-1)}{2}.$$

Applying a similar argument, one has $\sharp \{\Psi(ua) \mid ua \in \mathcal{F}_{\mathbf{t}_m}(n)\} = 1 + \frac{m(m-1)}{2}$.

Lemma 7. Let n = km + r with $k \ge 1$ and $1 \le r \le m - 1$. Then for every $a \in \Sigma_m$, we have

$$\sharp\{\Psi(au) \mid au \in \mathcal{F}_{\mathbf{t}_m}(n)\} = \sharp\{\Psi(ua) \mid ua \in \mathcal{F}_{\mathbf{t}_m}(n)\} = 1 + \frac{m(m-1)}{2}.$$

Proof. For $i \in [0, r]$ and $b \in \Sigma_m$, write

$$\mathcal{E}_{i,b} := \{ \alpha \sigma_m(v) \beta \in \mathcal{F}_{\mathbf{t}_m}(n) \mid \alpha = a |_i, \beta = b |_{r-i}, v \in \mathcal{F}_{\mathbf{t}_m}(k), a \triangleleft \alpha \sigma_m(v) \}$$

and $\mathcal{E}_i := \bigcup_{b \in \Sigma_m} \mathcal{E}_{i,b}$. By Lemma 5, $\mathcal{E}_{i,b} \neq \emptyset$ for every $b \in \Sigma_m$ and $i \in [0, m-1]$. Let $E_i = \bigcup_{b \in \Sigma_m} E_{i,b}$ where

$$E_{i,b} := \{\Psi(w) \mid w \in \mathcal{E}_{i,b}\} = \{\Psi(a|_i) + \Psi(b|_{r-i}) + k\mathbf{1}\}.$$

Moreover, for $i \in [0, r-1]$, $\sharp E_i = m$ and $\sharp E_r = 1$. Observe that $E_r \subset E_0$ and for $i \in [0, r-1]$,

$$\begin{cases} E_{i,b} = E_{0,a}, & \text{if } b \equiv a+i \pmod{m}; \\ E_{i,b} = E_{j,b}, & \text{if } b+r-i-a \equiv j \in [0,i-1]; \\ E_{i,b} \cap \left(\cup_{\ell=0}^{i-1} E_{\ell}\right) = \emptyset, & \text{otherwise.} \end{cases}$$

Thus, for $i \in [1, r-1]$, $\sharp \left(E_i \setminus \bigcup_{\ell=0}^{i-1} E_\ell \right) = m - i - 1$. So, $\sharp \left(\bigcup_{i=0}^r E_i \right) = 1 + \sum_{i=1}^r (m-i)$. For $i \in [1, m-r-1]$ and $b \in \Sigma_m$, write

$$\mathcal{D}_{i,b} := \{ \alpha \sigma_m(v) \beta \in \mathcal{F}_{\mathbf{t}_m}(n) \mid \alpha = a|_{m-i}, \beta = b|_{r+i}, v \in \mathcal{F}_{\mathbf{t}_m}(k-1) \}$$

and $\mathcal{D}_i := \bigcup_{b \in \Sigma_m} \mathcal{D}_{i,b}$. By Lemma 5, we have $\mathcal{D}_{i,b} \neq \emptyset$ for all $b \in \Sigma_m$ and $i \in [1, m - r - 1]$. Let $D_i = \bigcup_{b \in \Sigma_m} D_{i,b}$ where

$$D_{i,b} := \{\Psi(w) \mid w \in \mathcal{D}_{i,b}\} = \{\Psi(a|_{m-i}) + \Psi(b|_{r+i}) + (k-1)\mathbf{1}\}\$$

Note that $\sharp D_i = m$. Further, for $i \in [1, m - r - 1]$ and $j \in \Sigma_m$, letting $b \equiv a + m - i - j \pmod{m}$, we have

$$\begin{cases} D_{i,b} = E_{r-j,b}, & \text{if } j \in [0,r]; \\ D_{i,b} = D_{j-r,b}, & \text{if } j \in [r+1,r+i-1]; \\ D_{i,b} \cap \left[\left(\cup_{\ell=0}^{r} E_{\ell} \right) \cup \left(\cup_{s=1}^{i-1} D_{s} \right) \right] = \emptyset, & \text{if } j \in [r+i,m-1]. \end{cases}$$

Since $\{\Psi(au) \mid au \in \mathcal{F}_{\mathbf{t}_m}(n)\} = \left(\bigcup_{i=0}^r E_i\right) \cup \left(\bigcup_{i=1}^{m-r-1} D_i\right)$, we obtain that

$$\sharp \left[\left(\cup_{i=0}^{r} E_i \right) \cup \left(\cup_{i=1}^{m-r-1} D_i \right) \right] = 1 + \sum_{i=1}^{r} (m-i) + \sum_{i=r+1}^{m-1} (m-i) = 1 + \frac{m(m-1)}{2}.$$

Applying a similar argument, one has $\sharp \{\Psi(ua) \mid ua \in \mathcal{F}_{\mathbf{t}_m}(n)\} = 1 + \frac{m(m-1)}{2}$.

Lemma 8. For every $a, b \in \Sigma_m$ and $n \in \mathbb{N}$ with $n \ge m+1$, we have

$$\sharp\{\Psi(aub) \mid aub \in \mathcal{F}_{\mathbf{t}_m}(n)\} = \begin{cases} 1, & \text{if } n \equiv b - a + 1 \pmod{m}; \\ m, & \text{otherwise.} \end{cases}$$

Proof. Fix $a, b \in \Sigma_m$. Suppose n = km for some $k \ge 2$. Adopt the notations in the proof of Lemma 6. Recall that $\{w \in \mathcal{F}_{\mathbf{t}_m}(n) \mid a \triangleleft w\} = \bigcup_{i=0}^{m-1} \mathcal{E}_i$. Note that for $i \in [0, m-1]$, $\{w \in \mathcal{E}_i \mid b \triangleright w\} = \mathcal{E}_{i,b+i+1}$ and

$$E_{i,b+i+1} = \{\Psi(w) \mid w \in \mathcal{E}_{i,b+i+1}\} = \{\Psi(a|_i) + \Psi((b+i+1)|_{m-i}) + (k-1)\mathbf{1}\}.$$

If $b+i+1 \equiv a+i \pmod{m}$, then for all $i \in [0, m-1]$, $E_{i,b+i+1} = \{k1\}$. Consequently, $\sharp\{\Psi(aub) \mid i \in [0, m-1], i \in [0, m-1]\}$. $\begin{aligned} aub \in \mathcal{F}_{\mathbf{t}_m}(n) \} &= \sharp E_{0,a} = 1. \text{ If } b+i+1 \not\equiv a+i \pmod{m}, \text{ then for any } i, j \in [0, m-1] \text{ with } i \neq j, E_{i,b+i+1} \neq E_{j,b+j+1}. \text{ In this case, } \sharp \{\Psi(aub) \mid aub \in \mathcal{F}_{\mathbf{t}_m}(n)\} = \sharp \left(\cup_{i=0}^{m-1} E_{i,b+i+1} \right) = m. \end{aligned}$

Now suppose n = km + r where $k \ge 1$ and $1 \le r < m$. Adopt the notations in the proof of Lemma 7. Recall that $\{w \in \mathcal{F}_{\mathbf{t}_m}(n) \mid a \triangleleft w\} = (\bigcup_{i=0}^r \mathcal{E}_i) \cup (\bigcup_{i=1}^{m-r-1} \mathcal{D}_i)$. Again, for $i \in [0, r]$, $\{w \in \mathcal{E}_i \mid b \triangleright w\} = \mathcal{E}_{i,b-r+i+1}$ and for $j \in [1, m-r-1], \{w \in \mathcal{D}_i \mid b \triangleright w\} = \mathcal{D}_{i,b-r-i+1}$. Moreover,

$$\{\Psi(aub) \mid aub \in \mathcal{F}_{\mathbf{t}_m}(n)\} = \left(\cup_{i=0}^r E_{i,b-r+i+1}\right) \cup \left(\cup_{i=1}^{m-r-1} D_{i,b-r-i+1}\right).$$

If $b - a + 1 \equiv r \pmod{m}$, then for all $i \in [0, r]$ and $j \in [1, m - r - 1]$,

$$E_{i,b-r+i+1} = D_{j,b-r-j+1} = \{\Psi(a|_r) + k\mathbf{1}\}.$$

Therefore, $\sharp \{\Psi(aub) \mid aub \in \mathcal{F}_{\mathbf{t}_m}(n)\} = 1$. If $b - a + 1 \not\equiv r \pmod{m}$, then

 $E_{0,b-r+1}, E_{1,b-r+2}, \ldots, E_{r,b+1}, D_{1,b-r}, D_{2,b-r-1}, \ldots, D_{m-r-1,b-m+2}$ are pairwise disjoint. Thus $\sharp \{\Psi(aub) \mid aub \in \mathcal{F}_{\mathbf{t}_m}(n) \} = m.$

Recall that for $u \in \Sigma_m^*$, the Parikh vector and the extended Parikh vector of u are denoted by $\Psi(u)$ and $\Psi_k(u)$ respectively. The following two theorems give the accurate value of $b_{\mathbf{t}_m,2}(n)$ for every $n \ge m^2$.

Theorem 4. For every $k \ge m$, we have

$$b_{\mathbf{t}_m,2}(km) = b_{\mathbf{t}_m,1}(k) + m(m-1)[m(m-1)+1]$$

where $b_{\mathbf{t}_m,1}(\cdot)$ denotes the abelian complexity function of the infinite word \mathbf{t}_m .

Proof. Observe that $\mathcal{F}_{\mathbf{t}_m}(km) = F_1 \cup F_2$ with $F_1 \cap F_2 = \emptyset$ where

$$F_1 := \{ \sigma_m(v) \mid v \in \mathcal{F}_{\mathbf{t}_m}(k) \},$$

$$F_2 := \{ \alpha \sigma_m(v) \beta \in \mathcal{F}_{\mathbf{t}_m}(km) \mid (\alpha, \beta) \in \mathcal{S} \times \mathcal{P}, |\alpha| + |\beta| = m, v \in \mathcal{F}_{\mathbf{t}_m}(k-1) \}.$$

Applying Theorem 2, we have

$$b_{\mathbf{t}_m,2}(km) = \#\{\Psi_2(u) \mid u \in \mathcal{F}_{\mathbf{t}_m}(km)\} \\ = \#\{\Psi_2(u) \mid u \in F_1\} + \#\{\Psi_2(u) \mid u \in F_2\}.$$
(10)

Using Theorem 2 again, for $u = \sigma_m(v) \in F_1$ and $u' = \sigma_m(v') \in F_1$, $u \sim_2 u'$ if and only if $v \sim_1 v'$, i.e., $\Psi_2(u) = \Psi_2(u') \iff \Psi(v) = \Psi(v')$. Therefore,

$$\sharp\{\Psi_2(u) \mid u \in F_1\} = \sharp\{\Psi(v) \mid v \in \mathcal{F}_{\mathbf{t}_m}(k)\} = b_{\mathbf{t}_m,1}(k).$$
(11)

For $i \in [1, m - 1]$, write

$$F_{2,i,a,b} := \{ \alpha \sigma_m(v) \beta \in F_2 \mid \alpha = a \mid_i, b \triangleright \beta, v \in \mathcal{F}_{\mathbf{t}_m}(k-1) \}.$$

Then $F_2 = \bigcup_{i=1}^{m-1} \bigcup_{a,b \in \Sigma_m} F_{2,i,a,b}$. According to Theorem 2,

Using Theorem 2 again, if $(i, a, b) \neq (i', a', b')$, then for any $u \in F_{2,i,a,b}$ and $u' \in F_{2,i',a',b'}$, we have $u \nsim_2 u'$. So,

$$\sharp\{\Psi_2(u) \mid u \in F_2\} = \sum_{i=1}^{m-1} \sum_{a,b \in \Sigma_m} \sharp\{\Psi_2(w) \mid w \in F_{2,i,a,b}\}$$
$$= (m-1)m[1+(m-1)m].$$
(12)

Then the result follows from (10), (11) and (12).

Theorem 5. For every $n \ge m^2$ with $n \not\equiv 0 \pmod{m}$, we have

$$b_{\mathbf{t}_m,2}(n) = m^4 - 2m^3 + 2m^2.$$

Proof. Let n = km + r for some $k \ge m$ and $1 \le r \le m - 1$. Then the set $\mathcal{F}_{\mathbf{t}_m}(n)$ can be separated into four disjoint parts:

$$\begin{split} S_1 &:= \{ \alpha \sigma_m(v) \in \mathcal{F}_{\mathbf{t}_m}(n) \mid \alpha = a|_r, a \in \Sigma_m, v \in \mathcal{F}_{\mathbf{t}_m}(k) \}, \\ S_2 &:= \{ \sigma_m(v)\beta \in \mathcal{F}_{\mathbf{t}_m}(n) \mid \beta = b|_r, b \in \Sigma_m, v \in \mathcal{F}_{\mathbf{t}_m}(k) \}, \\ S_3 &:= \{ \alpha \sigma_m(v)\beta \in \mathcal{F}_{\mathbf{t}_m}(n) \mid \alpha = a|_i, \beta = b|_{r-i}, i \in [1, r-1], a, b \in \Sigma_m, v \in \mathcal{F}_{\mathbf{t}_m}(k) \}, \\ S_4 &:= \{ \alpha \sigma_m(v)\beta \in \mathcal{F}_{\mathbf{t}_m}(n) \mid \alpha = a|_i, \beta = b|_{m+r-i}, i \in [r+1, m-1], a, b \in \Sigma_m, v \in \mathcal{F}_{\mathbf{t}_m}(k-1) \}. \end{split}$$

According to Theorem 2,

$$\sharp\{\Psi_2(w) \mid w \in S_1\} = \sum_{a \in \Sigma_m} \sharp\{\Psi_2(a|_r \sigma_m(v)) \mid v \in \mathcal{F}_{\mathbf{t}_m}(k), a|_r \sigma_m(v) \in \mathcal{F}_{\mathbf{t}_m}(n)\}$$

$$= \sum_{a \in \Sigma_m} \sharp \{ \Psi(a'v) \mid a'v \in \mathcal{F}_{\mathbf{t}_m}(k+1), a' \equiv a+r \pmod{m} \}$$
$$= m \left(1 + \frac{m(m-1)}{2} \right), \tag{13}$$

where in the last step, we use Lemma 7. Similarly, we have

$$\sharp\{\Psi_2(w) \mid w \in S_2\} = m\left(1 + \frac{m(m-1)}{2}\right).$$
(14)

For $i \in [1, r-1]$, write

$$S_{3,i,a,b} := \left\{ \alpha \sigma_m(v) \beta \in \mathcal{F}_{\mathbf{t}_m}(n) \mid \alpha = a|_i, \beta = b|_{r-i}, v \in \mathcal{F}_{\mathbf{t}_m}(k) \right\}.$$

Then $S_3 = \bigcup_{i=1}^{r-1} \bigcup_{a,b \in \Sigma_m} S_{3,i,a,b}$. By Lemma 8,

$$\sharp\{\Psi(a'vb) \mid a'vb \in \mathcal{F}_{\mathbf{t}_m}(k+2), \ a' \equiv a+i \pmod{m}\} = \begin{cases} 1, & \text{if } b \equiv a+i+k+1 \pmod{m}; \\ m, & \text{otherwise.} \end{cases}$$

It follows from Theorem 2,

$$\sharp \{ \Psi_2(w) \mid w \in S_3 \} = \sum_{i=1}^{r-1} \sum_{a,b \in \Sigma_m} \sharp \{ \Psi(a'vb) \mid a'vb \in \mathcal{F}_{\mathbf{t}_m}(k+2), \ a' \equiv a+i \pmod{m} \}$$

= $(r-1)m[1+m(m-1)].$ (15)

For $i \in [r+1, m-1]$ and $a, b \in \Sigma_m$, write

$$S_{4,i,a,b} := \{ \alpha \sigma_m(v) \beta \in \mathcal{F}_{\mathbf{t}_m}(n) \mid \alpha = a \mid_i, \beta = b \mid_{m+r-i}, v \in \mathcal{F}_{\mathbf{t}_m}(k-1) \}.$$

By Theorem 2 and Lemma 8,

$$\sharp\{\Psi_2(w) \mid w \in S_4\} = \sum_{i=r+1}^{m-1} \sum_{a,b \in \Sigma_m} \sharp\{\Psi(a'vb) \mid a'vb \in \mathcal{F}_{\mathbf{t}_m}(k+1), \ a' \equiv a+i \pmod{m}\}$$
$$= (m-r-1)m[1+m(m-1)].$$
(16)

Combining (13), (14), (15), (16) and Theorem 2, we have

$$b_{\mathbf{t}_m,2}(n) = \#\{\Psi_2(w) \mid w \in \mathcal{F}_{\mathbf{t}_m}(n)\}$$

= $\sum_{j=1}^4 \#\{\Psi_2(w) \mid w \in S_j\} = m^4 - 2m^3 + 2m^2.$

Proof of Theorem 3. The result follows from Theorem 4 and Theorem 5 directly.

Acknowledgement

This work was supported by NSFC (Nos. 11801203, 11701202, 11871295), Guangzhou Science and Technology program (202102020294), Guangdong Basic and Applied Basic Research Foundation (2021A1515010056) and the Fundamental Research Funds for the Central Universities from SCUT (2020ZYGXZR041).

References

- J.-P. Allouche, J. Shallit, Automatic Sequences. Theory, Applications, Generalizations, Cambridge University Press, 2003.
- [2] J. Berstel, M. Crochemore, J.-É. Pin, Thue-Morse sequence and p-adic topology for the free monoid, Discrete Math. 76 (1989) 89-94.
- [3] J. Cassaigne, G. Richomme, K. Saari, L. Q. Zamboni, Avoiding Abelian powers in binary words with bounded Abelian complexity, Int. J. Found. Comput. Sci. 22(4) (2011) 905–920.
- [4] J. Chen, Z.-X. Wen, On the abelian complexity of generalized Thue-Morse sequences, Theor. Comput. Sci. 780 (2019) 66-73.
- [5] E. M. Coven, G. A. Hedlund, Sequences with minimal block growth, Math. Syst. Theory. 7 (1973) 138–153.
- [6] G. Fici, F. Mignosi, J. Shallit, Abelian-square-rich words, Theor. Comput. Sci. 684 (2017) 29-42
- [7] D. D. Freydenberger, P. Gawrychowski, J. Karhumäki, F. Manea, W. Rytter, *Testing k-binomial equivalence*, arXiv:1509.00622.
- [8] P. Karandikar, M. Kufleitner, Ph. Schnoebelen, On the index of Simon's congruence for piecewise testability, Inf. Process. Lett. 115 (2015) 515-519.
- [9] J. Karhumäki, A. Saarela, L. Q. Zamboni, On a generalization of Abelian equivalence and complexity of infinite words, J. Comb. Theory, Ser. A. 120 (2013) 2189-2206.
- [10] V. Keränen, Abelian squares are avoidable on 4 letters, in: ICALP92, in: Lecture Notes in Computer Science, vol. 623, Springer, 1992, pp. 41–52.
- [11] M. Lejeune, J. Leroy, M. Rigo, Computing the k-binomial complexity of the Thue-Morse word, J. Comb. Theory Ser. A. 176 (2020) 105284.
- [12] M. Lejeune, M. Rigo, M. Rosenfeld, Templates for the k-binomial complexity of the Tribonacci word, Adv. Appl. Math. 112 (2020) 101947.
- [13] J. Leroy, M. Rigo, M. Stipulanti, Generalized Pascal triangle for binomial coefficients of words, Adv. Appl. Math. 80 (2016) 24-47.
- [14] X.-T. Lü, J. Chen, Z.-X. Wen, W. Wu, On the abelian complexity of the Rudin-Shapiro sequence, J. Math. Anal. Appl. 451 (2017) 822–838.
- [15] B. Madill, N. Rampersad, The abelian complexity of the paperfolding word, Discrete Math. 313(7) (2013) 831–838.
- [16] M. Morse, G. A. Hedlund, Symbolic dynamics II, Amer. J. Math. 62 (1938) 815-866.
- [17] J. Peltomäki, M. A. Whiteland, Every nonnegative real number is an Abelian critical exponent, in: WORDS 2019, in: Lecture Notes in Computer Science, vol. 11682, Springer, 2019, pp. 275–285.
- [18] J. Peltomäki, M. A. Whiteland, Avoiding Abelian powers cyclically, Adv. Appl. Math. 121 (2020) 102095.

- [19] J.-É. Pin, P.V. Silva, A noncommutative extension of Mahler's theorem on interpolation series, Eur. J. Comb. 36 (2014) 564-578.
- [20] S. Puzynina, Abelian properties of words, in: WORDS 2019, in: Lecture Notes in Computer Science, vol. 11682, Springer, 2019, pp. 28–45.
- [21] S. Puzynina, L. Q. Zamboni, Abelian returns in Sturmian words, J. Comb. Theory, Ser. A 120(2) (2013) 390–408
- [22] M. Rao, M. Rigo, P. Salimov, Avoiding 2-binomial squares and cubes, Theor. Comput. Sci. 572 (2015) 83-91.
- [23] G. Richomme, K. Saari, L. Q. Zamboni, Abelian complexity in minimal subshifts, J. Lond. Math. Soc. 83 (2011) 79–95.
- [24] M. Rigo, *Relations on words*, Indag. Math. (N.S.) 28 (2017) 183-204.
- [25] M. Rigo, P. Salimov, Another generalization of abelian complexity : binomial complexity of infinite words, Theor. Comput. Sci. 601 (2015) 47-57.