Noname manuscript No.
(will be inserted by the editor)

Disjunctivity and Other Properties of Sets of Pseudo-Bordered Words

Lila Kari · Manasi S. Kulkarni

Received: date / Accepted: date

Abstract The concepts of pseudo-bordered and pseudo-unbordered words are in large part motivated by research in theoretical DNA computing, wherein the Watson-Crick complementarity of DNA strands is modelled as an antimorphic involution, that is, a function θ which is an antimorphism, $\theta(uv) = \theta(v)\theta(u)$, and an involution, $\theta(\theta(u)) = u$, for all words u, v over the DNA alphabet. In particular, a word w is said to be θ-bordered (or pseudo-bordered) if there exists a word $v \in \Sigma^+$ that is a proper prefix of w, while $\theta(v)$ is a proper suffix of w. A word which is not θ-bordered is θ-unbordered. This paper continues the exploration of properties (for the case where θ is a morphic involution) of the set of θ-unbordered words, $D_\theta(1)$, and the sets of words that have exactly i θ-borders, $D_\theta(i)$, $i \geq 2$. We prove that, under some conditions, the set $D_\theta(i)$ is disjunctive for all $i \geq 1$, and that the set $D^i_\theta(1) \setminus D(i)$ is disjunctive for all $i \geq 2$, where $D(i)$ denotes the set of words with exactly i borders. We also discuss conditions for catenations of languages of θ-unbordered words to remain θ-unbordered, and anticipate further generalizations by showing that the set of all θ-bordered words is not context-free for all morphisms θ over an alphabet Σ with $|\Sigma| \geq 3$ such that $\theta(a) \neq a$ for all $a \in \Sigma$ and θ^3 equals the identity function on Σ.

Keywords Bordered words · unbordered words · pseudo-bordered words · pseudo-unbordered words · DNA computing · disjunctivity

1 Introduction

Combinatorics on words, coding theory, and formal language theory have had a wide range of applications ranging from bioinformatics, to cryptography, to DNA computing. For example, the concepts of periodicity and primitivity are at the root of pattern-matching and data compression algorithms, [5,6,33], and the study of

This research was supported by Natural Science and Engineering Council of Canada (NSERC) Discovery Grant R2824A01 to L.K.

Dept. of Computer Science, University of Western Ontario, London, ON, N6A 5B7, Canada, E-mail: lila.kari@uwo.ca, mkulkar3@uwo.ca

codes is essential in determining the unique decipherability of encoded messages, [28]. Notably, the recent connection with DNA computing has motivated a new line of study wherein classical concepts are generalized to ones where the identity function is replaced with more general pseudo-identity functions. A representative example of such a generalization is the concept of antimorphic involution which models the DNA Watson-Crick complementarity, as described below.

DNA single strands can be viewed as strings over the DNA alphabet $\{A, C, G, T\}$. The Watson-Crick complementarity is the property whereby two DNA *single* strands of opposite orientation and with complementary "letters" at each position can bind together by hydrogen bonds to form a DNA *double* strand with its well-known double helical structure [29]. Given an alphabet Σ, an antimorphic involution θ is a function that is an antimorphism, that is, $\theta(uv) = \theta(v)\theta(u)$, $\forall u, v \in \Sigma^*$, and an involution, that is, $\theta(\theta(x)) = x$, $\forall x \in \Sigma^*$. Thus, the first property (antimorphism) models the fact that DNA single strands that bind to each other must have opposite orientations, and the second property (involution) models the letter-to-letter complementarity of the two single strands (whereby A binds to a T, and C binds to a G) that is necessary for the binding to occur.

Note that a DNA single strand and its Watson-Crick complement are informationally equivalent, since one uniquely determines the other and viceversa. Thus, a DNA strand and its Watson-Crick complement can be viewed in a sense as "identical", and this motivated the idea of generalizing the notion of identity function to pseudo-identity functions, such as antimorphic involutions. Some of the new concepts in combinatorics on words and coding theory that were thus obtained are: Pseudo-periodicity, [8, 23], pseudo-commutativity, pseudo-conjugacy, [20], pseudo-palindrome, [21, 9], involution codes, [3, 16, 17], etc. Some of these concepts were further generalized in [10–12] by replacing the morphic involution with length-preserving, erasing and uniform morphism functions. Also, independently, the notion of periodicity was extended to periodic-like words, [2], weakly periodic words, [7], also known as Abelian periodic words, [4], and pseudoperiodic words, [1].

A non-empty word w is said to be bordered if there exists a word that is a proper prefix and a proper suffix of w. A word which is not bordered is called unbordered. In [19] the notion bordered word was generalized to that of a θ-bordered word (also called pseudo-bordered word), where θ is (anti)morphic involution: A word w is said to be θ-bordered if there exists a word $v \in \Sigma^+$ that is a proper prefix of w, while $\theta(v)$ is a proper suffix of w. Naturally, a word which is not θ-bordered is θ-unbordered. Properties of θ-bordered and θ-unbordered words were explored in, e.g., [15, 19]. The classical notions of bordered and unbordered words have also been generalized to pseudo-knot-bordered words in [22], where a non-empty word w is said to be pseudo-knot-bordered if $w = xy\alpha = \beta\theta(yx)$ for $\alpha, \beta, x, y \in \Sigma^+$.

In this paper we continue to explore the properties of θ-bordered and θ-unbordered words, for *morphic involutions* θ. The main focus is on disjunctivity properties of sets of θ-bordered words and some other related languages. The paper is organized as follows. Section 2 includes basics definitions and notions used throughout the paper. In Section 3 we prove, e.g., that under some conditions, the set of all θ-bordered words with exactly i θ-borders, $D_\theta(i)$, is disjunctive for all $i \geq 1$ (Theorem 3). In Section 4 and 5, we discuss relationships between and among the sets $D_\theta(1)$, the set of all θ-unbordered words, and the set $D(i)$, of all bordered words with exactly i borders. In particular, we show that, under some conditions, the set $D_\theta^i(1) \setminus D(i)$ is disjunctive for all $i \geq 2$ (Theorem 4). In Section 6 we discuss some conditions for catenations of languages of θ-unbordered words to remain θ-unbordered, and offer a preview of further generalizations of these results by proving that the set of all θ-bordered words is not context-free for all morphisms θ over an alphabet Σ with $|\Sigma| \geq 3$ such that $\theta(a) \neq a$ for all $a \in \Sigma$ and θ^3 equals the identity function on Σ.

2 Basic definitions and notations

An alphabet Σ is a finite non-empty set of symbols. Σ^* denotes the set of all words over Σ, including the empty word λ. Σ^+ is the set of all non-empty words over Σ. The length of a word $u \in \Sigma^*$ (i.e. the number of symbols in a word) is denoted by $|u|$. By Σ^m we denote the set of all words of length $m > 0$ over Σ. The complement of a language $L \subseteq \Sigma^*$ is $L^c = \Sigma^* \setminus L$. For a language $L \subseteq \Sigma^*$ and $i \geq 2$, let $L^{(i)} = \{u^i | u \in L\}$ and $L^1 = L$ and $L^n = L^{n-1}L$ for $n \geq 2$. A word is called *primitive* if it cannot be expressed as a power of another word. Let Q denote the set of all primitive words. A function $\theta : \Sigma^* \to \Sigma^*$ is said to be a *morphism* if for all words $u, v \in \Sigma^*$ we have that $\theta(uv) = \theta(u)\theta(v)$, an *antimorphism* if $\theta(uv) = \theta(v)\theta(u)$, and an *involution* if θ^2 is an identity on Σ^*. If for all $a \in \Sigma$, $|\theta(a)| = 1$, then θ is called *literal* (anti)morphism[1]. A θ*-power* of a word u, [8] is a word of the form $u_1u_2 \ldots u_n$ for $n \geq 1$ where $u_1 = u$ and $u_i \in \{u, \theta(u)\}$ for $2 \leq i \leq n$. A word is called θ*-primitive*, [8], if it cannot be expressed as a θ-power of another word. Let Q_θ denote the set of all θ-primitive words. For (anti)morphic involution θ, a word $u \in \Sigma^*$ is called a θ*-palindrome*, [21, 9], if $u = \theta(u)$. Let P_θ denote the set of all θ-palindromes.

For a language $L \subseteq \Sigma^*$, the *principal congruence* P_L determined by L is defined as follows: for any $x, y \in \Sigma^*$ such that $x \neq y$, $x \equiv y(P_L)$ if and only if $uxv \in L \Leftrightarrow uyv \in L$ for all $u, v \in \Sigma^*$. The index of P_L is the number of equivalence classes of P_L. L is said to be *disjunctive* if P_L is the identity, i.e., for any $x \neq y \in \Sigma^*$ there exists $u, v \in \Sigma^*$ such that $uxv \in L$ and $uyv \notin L$ or viceversa. A language $L \subseteq \Sigma^*$ is said to be *dense* if for all $u \in \Sigma^*$, $L \cap \Sigma^* u \Sigma^* \neq \emptyset$. Every disjunctive language is dense and every dense language contains a disjunctive subset, [27].

Definition 1 1. For $v, w \in \Sigma^*$, w is a prefix of v ($w \leq_p v$) iff $v \in w\Sigma^*$.
2. For $v, w \in \Sigma^*$, w is a suffix of v ($w \leq_s v$) iff $v \in \Sigma^* w$.
3. $\leq_d = \leq_p \cap \leq_s$.
4. For $u \in \Sigma^*$, $v \in \Sigma^*$ is said to be a border of u if $v \leq_d u$, i.e., $u = vx = yv$.
5. For $v, w \in \Sigma^*$, w is a proper prefix of v ($w <_p v$) iff $v \in w\Sigma^+$.
6. For $v, w \in \Sigma^*$, w is a proper suffix of v ($w <_s v$) iff $v \in \Sigma^+ w$.
7. $<_d = <_p \cap <_s$.
8. For $u \in \Sigma^*$, $v \in \Sigma^*$ is said to be a proper border of u if $v <_d u$.
9. For $u \in \Sigma^+$, denote by $L_d(u) = \{v \in \Sigma^* | v <_d u\}$, the set of all borders of a word $u \in \Sigma^*$.
10. $\nu_d(u) = |L_d(u)|$.
11. Denote by $D(i) = \{u \in \Sigma^+ | \nu_d(u) = i\}$, the set of all words with exactly i borders for $i \geq 1$.

[1] By (anti)morphism we mean either a morphism or an antimorphism.

12. A word $u \in \Sigma^+$ is said to be a bordered word if there exists $v \in \Sigma^+$ such that $v <_d u$, i.e., $u = vx = yv$ for some $x, y \in \Sigma^+$.
13. A non-empty word which is not bordered is called unbordered. Thus, $D(1)$ is the set of all unbordered words over Σ.

For a word w, $\mathrm{Pref}(w) = \{u \in \Sigma^+ | \exists v \in \Sigma^*, w = uv\}$ and $\mathrm{Suff}(w) = \{u \in \Sigma^+ | \exists v \in \Sigma^*, w = vu\}$ denotes the set of all prefixes and suffixes respectively. Similarly, the set of all proper prefixes and proper suffixes of a word w can be defined as $\mathrm{PPref}(w) = \{u \in \Sigma^+ | \exists v \in \Sigma^+, w = uv\}$ and $\mathrm{PSuff}(w) = \{u \in \Sigma^+ | \exists v \in \Sigma^+, w = vu\}$ respectively. For further notions in formal language theory and combinatorics on words the reader is referred to [13,25,27,32].

The following definitions extend the notion of bordered and unbordered words to θ-bordered and θ-unbordered words and for any (anti)morphism on Σ^*.

Definition 2 [19] Let θ be either a morphism or an antimorphism on Σ^*.

1. For $v, w \in \Sigma^*$, w is a θ-prefix of v ($w \leq_p^\theta v$) iff $v \in \theta(w)\Sigma^*$.
2. For $v, w \in \Sigma^*$, w is a θ-suffix of v ($w \leq_s^\theta v$) iff $v \in \Sigma^*\theta(w)$.
3. $\leq_d^\theta = \leq_p \cap \leq_s^\theta$.
4. For $u \in \Sigma^*$, $v \in \Sigma^*$ is said to be a θ-border of u if $v \leq_d^\theta u$, i.e., $u = vx = y\theta(v)$.
5. For $w, v \in \Sigma^*$, w is a proper θ-prefix of v ($w <_p^\theta v$) iff $v \in \theta(w)\Sigma^+$.
6. For $w, v \in \Sigma^*$, w is a proper θ-suffix of v ($w <_s^\theta v$) iff $v \in \Sigma^+\theta(w)$.
7. $<_d^\theta = <_p \cap <_s^\theta$.
8. For $u \in \Sigma^*$, $v \in \Sigma^*$ is said to be a proper θ-border of u if $v <_d^\theta u$.
9. For $u \in \Sigma^+$, define by $L_d^\theta(u) = \{v \in \Sigma^* | v <_d^\theta u\}$, the set of all θ-borders of a word $u \in \Sigma^*$.
10. $\nu_d^\theta(u) = |L_d^\theta(u)|$.
11. Denote by $D_\theta(i) = \{u \in \Sigma^+ | \nu_d^\theta(u) = i\}$, the set of all words with exactly i θ-borders for $i \geq 1$.
12. A word $u \in \Sigma^+$ is said to be θ-bordered if there exists $v \in \Sigma^+$ such that $v <_d^\theta u$, i.e., $u = vx = y\theta(v)$ for some $x, y \in \Sigma^+$.
13. A nonempty word which is not θ-bordered is called θ-unbordered. Thus, $D_\theta(1)$ is the set of all θ-unbordered words over Σ.

Recall that every disjunctive language has infinitely many principle congruence classes whereas the number of principle congruence classes for regular language is finite. Hence, it is clear that disjunctive languages are not regular.

The following proposition provides a necessary and sufficient condition for a language to be disjunctive, and will be used throughout this paper.

Proposition 1 *[27] Let $L \subseteq \Sigma^*$. Then the following two statements are equivalent:*

1. *L is a disjunctive language.*
2. *If $u, v \in \Sigma^+$, $u \neq v$, $|u| = |v|$, then $u \not\equiv v(P_L)$.*

While proving disjunctivity or any other properties of the sets of words with exactly i borders or i θ-borders, $D(i)$ or $D_\theta(i)$ respectively, one of the important tools is the knowledge about the number of borders and θ-borders of a word. Proposition 2 characterizes the number of borders of a power of a primitive word.

Proposition 2 *[14] For any $f \in Q$ and $j \geq 1$, $\nu_d(f^j) = \nu_d(f) + j - 1$.*

Similarly, Lemma 1 provides a characterization for the number of θ-borders of a θ-palindrome, for morphic involutions.

Lemma 1 *[19] Let u be a θ-palindromic primitive word and j be an integer, $j > 1$. Then, for a morphic involution θ, $\nu_d^\theta(u^j) = \nu_d^\theta(u) + j - 1$.*

The following lemma provides a sufficient condition for a word to be bordered.

Lemma 2 *[14] Let $u \in \Sigma^+ \backslash D(1)$. Then there exists $v \in \Sigma^*$ with $|v| \leq \frac{|u|}{2}$ such that $v <_d u$.*

By the definition of an unbordered word, it is clear that the set of all unbordered words $D(1)$ is a subset of set of all primitive words Q, i.e., $D(1) \subseteq Q$. A similar inclusion does not hold in the case of set of all θ-unbordered words $D_\theta(1)$ and the set of all θ-primitive words Q_θ for a morphic involution θ, as demonstrated by following example. The example also demonstrates the fact that Q_θ is not a subset of $D_\theta(1)$.

Example 1 Let $\Sigma = \{a, b, c\}$, θ be a morphic involution such that $\theta(a) = b$, $\theta(b) = a$ and $\theta(c) = c$. Let $u = abaa$, then $u \in D_\theta(1)$ but $u = abaa = a\theta(a)aa \notin Q_\theta$ and hence $D_\theta(1) \nsubseteq Q_\theta$. Now, let $v = acb$, then $u \in Q_\theta$ but $u = acb = ac\theta(a) \notin D_\theta(1)$ and hence $Q_\theta \nsubseteq D_\theta(1)$.

However, for a morphic involution θ, the set $D_\theta(1) \cap Q_\theta \neq \emptyset$. For example, if $\Sigma = \{a, b, c\}$ such that $\theta(a) = b$, $\theta(b) = a$ and $\theta(c) = c$, then $abc \in D_\theta(1) \cap Q_\theta$. Moreover, the set $D_\theta(i) \cap Q_\theta \neq \emptyset$ for all $i \geq 1$.

3 Disjunctivity properties of $D_\theta(i)$

In [14] it was shown that the languages $D(i)$, $D(i) \cap Q$ and $D(i) \cap Q^{(j)}$ are disjunctive for $i \geq j \geq 1$. In this section, we will prove the disjunctivity of the set $D_\theta(i)$ for all $i \geq 1$ (Theorem 3). Also, we know from Example 1 that neither $D_\theta(1) \subseteq Q_\theta$ nor $Q_\theta \subseteq D_\theta(1)$ but $D_\theta(i) \cap Q_\theta \neq \emptyset$ for all $i \geq 1$. Furthermore, in this section we will prove that the set $D_\theta(i) \cap Q_\theta^{2i-2}$ is disjunctive for $i \geq 2$ (Corollary 1).

In the previous section, we have seen a sufficient condition for a word to be bordered. The following lemma provides a sufficient condition for a word to be θ-bordered in the case when θ is a morphic involution.

Lemma 3 *[18] Let θ be a morphic involution and let $u \in \Sigma^+ \backslash D_\theta(1)$. Then there exists $v \in \Sigma^*$ with $|v| \leq \frac{|u|}{2}$ such that $v <_d^\theta u$.*

Theorem 1 and 2 are mentioned for completeness.

Theorem 1 *[18] Let θ be a morphic involution on Σ^*, where Σ is an alphabet with $|\Sigma| \geq 2$ that contains letters $a \neq b$ such that $a \neq \theta(b)$. Then the set of all θ-unbordered words, $D_\theta(1)$ and set of words with exactly two θ-borders $D_\theta(2)$ are disjunctive.*

Theorem 2 *[18] Let θ be a morphic involution on Σ^*, where Σ is an alphabet with $|\Sigma| \geq 3$ that contains letters $a \neq b$ such that $a \notin \{\theta(b), \theta(a)\}$. Then the set $[D_\theta(1) \cap D(1)]^n$ is disjunctive for any even number $n \geq 2$.*

While Theorem 1 proves the disjunctivity of the set $D_\theta(i)$ for the cases $i = 1, 2$, we will prove (Theorem 3) that the set $D_\theta(i)$ is disjunctive for all $i \geq 3$ as well.

We first need several auxiliary results. In the previous section, we mentioned a characterization of the number of borders of a power of a primitive word. Now, we will provide a characterization of the number of θ-borders of a θ-power of a θ-unbordered word for morphic involution θ (Proposition 3). Note that here we consider a special case of a θ-power of a word $w = u_1u_2 \ldots u_n$, where $u_i = u$ when i is odd and $u_i = \theta(u)$ when i is even for $1 \leq i \leq n$. The following lemma is needed for the proof of Proposition 3.

Lemma 4 *Let θ be morphic involution such that $\theta(a) \neq a$ for all $a \in \Sigma$. If $u \in D_\theta(1)$, then for $w = (u\theta(u))^k$, $u' <_p u$ we have that $(u\theta(u))^j u', (u\theta(u))^j u\theta(u') \notin L_d^\theta(w)$ for all $k > j \geq 1$.*

Proof We will prove the result by contradiction. Let $k > j \geq 1$ and $u' <_p u$.

First, assume that $(u\theta(u))^j u' <_d^\theta w$. Then, there exists $\alpha, \beta \in \Sigma^+$ such that $w = (u\theta(u))^k = (u\theta(u))^j u'\alpha = \beta(\theta(u)u)^j\theta(u')$. Since $|u'| < |u|$, we have that $\theta(u') <_s \theta(u)$ which implies $\theta(u) = u''\theta(u')$ for $u'' \in \Sigma^+$. This implies that $\theta(u'') <_p u$ since $u'' <_p \theta(u)$. But then, $(u\theta(u))^k = (u\theta(u))^{k-1}uu''\theta(u') = \beta(\theta(u)u)^{j-1}\theta(u)u\theta(u')$ which implies $u'' <_s u$ since $|u''| < |u|$ which further implies that $\theta(u'') <_d^\theta u$, i.e., $u \notin D_\theta(1)$, a contradiction. Hence, $(u\theta(u))^j u' \notin L_d^\theta(w)$.

Now, let $(u\theta(u))^j u\theta(u') <_d^\theta w$. Then there exists $\alpha', \beta' \in \Sigma^+$ such that $w = (u\theta(u))^k = (u\theta(u))^j u\theta(u')\alpha' = \beta'(\theta(u)u)^j\theta(u)u'$. Since $|u'| < |u|$, which implies $u' <_s \theta(u)$, i.e., $\theta(u') <_s u$ which further implies $u' <_d^\theta u$ and hence $u \notin D_\theta(1)$, a contradiction. □

Proposition 3 *Let θ be morphic involution such that $\theta(a) \neq a$ for all $a \in \Sigma$. If $u \in D_\theta(1)$, then $w = (u\theta(u))^n \in D_\theta(n+1)$ for all $n \geq 1$.*

Proof We will prove this statement by induction on n.

Let $n = 1$, then for $w = u\theta(u)$, since $u \in D_\theta(1)$, $L_d^\theta(w) = \{\lambda, u\}$. Hence $w = u\theta(u) \in D_\theta(2)$.

Let $n = 2$, then for $w = u\theta(u)u\theta(u)$, by Lemma 4 $u\theta(u)u', u\theta(u)u\theta(u') \notin L_d^\theta(w)$ where $u' \in \text{PPref}(u)$ and hence $L_d^\theta(w) = \{\lambda, u, u\theta(u)u\}$. Thus $w \in D_\theta(3)$.

Let us assume that the result holds for $n = k$, i.e., $w = (u\theta(u))^k \in D_\theta(k+1)$.

Now, we will prove that the result holds for $n = k+1$. We have $w = (u\theta(u))^{k+1} = (u\theta(u))^k u\theta(u)$. By inductive hypothesis, we know that $(u\theta(u))^k \in D_\theta(k+1)$. Also, by Lemma 4, $(u\theta(u))^k u', (u\theta(u))^k u\theta(u') \notin L_d^\theta(w)$ for some $u' <_p u$. Thus, $L_d^\theta(w) = L_d^\theta((u\theta(u))^k) \cup \{(u\theta(u))^k u\}$ and hence $w \in D_\theta(k+2)$.

Hence, $w = (u\theta(u))^n \in D_\theta(n+1)$ for all $n \geq 1$. □

In the preceding two results we considered a special case of θ-powers, namely, words w consisting of alternations of u and $\theta(u)$. Under certain conditions, if in such words the first occurrence of u is replaced by $v \neq u$, then the word w becomes θ-unbordered, as showed by the following result.

Lemma 5 *Let θ be a morphic involution on Σ^*, where Σ is an alphabet with $|\Sigma| > 2$ that contains letters $a \neq b$ such that $a \neq \theta(b)$. Let $x \neq y$, $x, y \in \Sigma^m$, $m > 0$, $x, y \in \theta(b)\Sigma^*$. Then, for all $i \geq 2$,*

$$a^m y\theta(b)(\theta(a^m x\theta(b))a^m x\theta(b))^{i-2}\theta(a^m x\theta(b)) \in D_\theta(1).$$

Proof Let us assume that

$$w = a^m y\theta(b)(\theta(a^m x\theta(b))a^m x\theta(b))^{i-2}\theta(a^m x\theta(b)) \notin D_\theta(1).$$

Then there exists $v \in \Sigma^+$ such that $v <_d^\theta w$, i.e., $w = v\alpha = \beta\theta(v)$ for some $\alpha, \beta \in \Sigma^+$. Let $w = w'\theta(a^m x\theta(b))$ where $w' = a^m y\theta(b)(\theta(a^m x\theta(b))a^m x\theta(b))^{i-2}$. Then, by Lemma 3, it is enough to consider only the cases when $1 \leq |v| < (2m+1)(i-1)$.

Case 1: $v = a^k$ for $1 \leq k \leq m$. Then,

$$w = w'\theta(a^m x\theta(b)) = \beta\theta(a^k)$$

which implies $\theta(a) = b$, a contradiction.

Case 2: $v = a^m y'$ for $y = y'y''$ where $y' \in \Sigma^+$ and $y'' \in \Sigma^*$. Then,

$$w = w'\theta(a^m x\theta(b)) = \beta\theta(a^m y').$$

Now, since $|a^m y'| < |a^m x\theta(b)|$, $\theta(a^m y') <_s \theta(a^m x\theta(b))$, i.e., $\theta(a^m x\theta(b)) = \theta(a^m)b\theta(x')b = \beta'\theta(a^m y')$ for $x = \theta(b)x'$ where $x' \in \Sigma^*$. This implies $\theta(a) = b$, a contradiction.

Case 3: $v = a^m y\theta(b)$. Then,

$$w = w'\theta(a^m x\theta(b)) = \beta\theta(a^m y\theta(b))$$

which implies $x = y$, a contradiction.

Case 4: $v = a^m y\theta(b)\theta(a^k)$ for $1 \leq k \leq m$. Then,

$$w = w'\theta(a^m x\theta(b)) = \beta\theta(a^m y\theta(b))a^k$$

which implies $a = b$, a contradiction.

Case 5: $v = a^m y\theta(b)\theta(a^m x_1)$ for $x = x_1x_2$ where $x_1 \in \Sigma^+$ and $x_2 \in \Sigma^*$. Then,

$$w = w'\theta(a^m x\theta(b)) = \beta\theta(a^m y\theta(b))a^m x_1.$$

Now, since $2m \geq |a^m x_1| \geq |\theta(x)b| = m+1$, $\theta(x)b \leq_s a^m x_1$, i.e., $a^m x_1 = \alpha'\theta(x)b = \alpha' b\theta(x')b$ with $|\alpha'| < m$ and $x = \theta(b)x'$ for $x' \in \Sigma^*$. This implies $a = b$, a contradiction.

Case 6: $v = a^m y\theta(b)\theta(a^m x\theta(b))$. Then,

$$w = w'\theta(a^m x\theta(b)) = \beta\theta(a^m y\theta(b))a^m x\theta(b)$$

which implies $\theta(a) = a$, a contradiction.

Case 7: $v = a^m y\theta(b)\theta(a^m x\theta(b))a^k$ for $1 \leq k \leq m$. Then,

$$w = w'\theta(a^m x\theta(b)) = \beta\theta(a^m y\theta(b))a^m x\theta(b)\theta(a^k).$$

which implies $\theta(a) = b$, a contradiction

Case 8: $v = a^m y\theta(b)\theta(a^m x\theta(b))a^m x_1'$ for $x = x_1'x_2'$ where $x_1' \in \Sigma^+$ and $x_2' \in \Sigma^*$. Then,

$$w = w'\theta(a^m x\theta(b)) = \beta\theta(a^m y\theta(b))a^m x\theta(b)\theta(a^m x_1').$$

Now, since $|ba^m x_1'| \leq |a^m x\theta(b)| = 2m+1$, $\theta(ba^m x_1') \leq_s \theta(a^m x\theta(b))$, i.e., $\theta(a^m x\theta(b)) = \alpha_2\theta(ba^m x_1')$ where $|\alpha_2| < m$. This implies $\theta(a) = \theta(b)$, i.e., $a = b$, a contradiction.

Case 9: $v = a^m y\theta(b)(\theta(a^m x\theta(b))a^m x\theta(b))^k$ where $1 \le k < \frac{i-2}{2}$. Then,

$$\begin{aligned} w &= a^m y\theta(b)(\theta(a^m x\theta(b))a^m x\theta(b))^{i-2}\theta(a^m x\theta(b)) \\ &= a^m y\theta(b)(\theta(a^m x\theta(b))a^m x\theta(b))^{i-2-k}\theta(a^m x\theta(b))(a^m x\theta(b)\theta(a^m x\theta(b)))^k \\ &= \beta\theta(a^m y\theta(b))(a^m x\theta(b)\theta(a^m x\theta(b)))^k. \end{aligned}$$

which implies $\theta(x) = \theta(y)$, i.e., $x = y$, a contradiction.

Since all the cases lead to a contradiction, $w \in D_\theta(1)$. □

The next lemma is used for proving the main result of this section.

Lemma 6 *[18] Let θ be a morphic involution and $a, b \in \Sigma$, $a \neq b$. Let $x, y \in \Sigma^m$, $m > 0$. Then*

1. $a^m x\theta(b) \in D_\theta(1)$.
2. *If $a \neq \theta(a)$, $x = \theta(b)x'$, $x' \in \Sigma^*$ and $k > m$, then $(a^k y\theta(b))(a^k x\theta(b)) \in D_\theta(1)$.*

Now, we will prove the main result of the section which shows that, under certain conditions, the set of words with exactly i θ-borders, $D_\theta(i)$, is disjunctive for all $i \ge 1$ and morphic involutions θ.

Theorem 3 *Let θ be a morphic involution on Σ^*, where Σ is an alphabet with $|\Sigma| > 2$ that contains letters $a \neq b$ such that $a \neq \theta(b)$ and $\theta(a) \neq a$ for all $a \in \Sigma$. Then the set of all θ-bordered words with exactly i θ-borders, $D_\theta(i)$, is disjunctive for all $i \ge 1$.*

Proof By Theorem 1, $D_\theta(i)$ is disjunctive for $i = 1, 2$. Now, we will prove the result for $i \ge 3$. Let $x, y \in \Sigma^n$, $x \neq y$, $m = n + 1$, $n > 0$. Let $u = a^m\theta(b)$, $v = \theta(b)(\theta(a^m\theta(b)x\theta(b))a^m\theta(b)x\theta(b))^{i-2}\theta(a^m\theta(b)x\theta(b))$. Since $a \neq b$, by Lemma 6, we have $a^m\theta(b)x\theta(b) \in D_\theta(1)$ and by Proposition 3,

$$uxv = [a^m\theta(b)x\theta(b)\theta(a^m\theta(b)x\theta(b))]^{i-1} \in D_\theta(i).$$

Further by Lemma 5,

$$uyv = a^m\theta(b)y\theta(b)[\theta(a^m\theta(b)x\theta(b))a^m\theta(b)x\theta(b)]^{i-2}\theta(a^m\theta(b)x\theta(b)) \in D_\theta(1).$$

Therefore, $x \not\equiv y(P_{D_\theta(i)})$ for every $x, y \in \Sigma^+$, $x \neq y$, $|x| = |y|$ and $i \ge 3$. Hence, by Proposition 1, $D_\theta(i)$ is disjunctive for $i \ge 1$. □

Let $\{a, b\} \subseteq \Sigma$ be such that $a \notin \{b, \theta(b)\}$ and θ be a morphic involution. Then for $x \in \Sigma^n$, $n > 0$ and $m = n+1$, it is clear that $a^m\theta(b)x\theta(b), \theta(a^m\theta(b)x\theta(b)) \in Q_\theta$. Thus, we have following result as a consequence of Theorem 3.

Corollary 1 *Let θ be a morphic involution on Σ^*, where Σ is an alphabet with $|\Sigma| > 2$ that contains letters $a \neq b$ such that $a \neq \theta(b)$ and $\theta(a) \neq a$ for all $a \in \Sigma$. Then the set $D_\theta(i) \cap Q_\theta^{2i-2}$ is disjunctive for all $i \ge 2$.*

4 Disjunctivity of the set $D_\theta^i(1)\backslash D(i)$

Let us consider the relationship between the set of all words with exactly i θ-borders, $D_\theta(i)$, and the set of all words with exactly i borders, $D(i)$, for $i \geq 1$, an alphabet Σ, and morphic involutions θ such that $\theta(a) \neq a$ for all $a \in \Sigma$. It is clear that in general neither $D_\theta(i) \subseteq D(i)$ nor $D(i) \subseteq D_\theta(i)$. However, the set $D_\theta(i) \cap D(i) \neq \emptyset$ for all $i \geq 1$. For example, if $\Sigma = \{a, b, c\}$ such that $\theta(a) = b$, $\theta(b) = a$ and $\theta(c) = c$, then $abc \in D_\theta(1) \cap D(1)$ and $abba \in D_\theta(2) \cap D(2)$. Moreover, Theorem 2 proved that the set $(D_\theta(1) \cap D(1))^n$ is disjunctive for any even number $n \geq 2$. In this section, we will show that, under certain conditions, the set $D_\theta^i(1)\backslash D(i)$ is disjunctive for $i \geq 2$ (Theorem 4).

In order to show that the language $D_\theta^i(1)\backslash D(i)$ is disjunctive, we need to characterize some catenations of unbordered and θ-unbordered words. The following proposition shows such a relationship.

Proposition 4 *[31] Let $\{a, b\} \in \Sigma$ and let $x, y \in b\Sigma^*$ with $x \neq y$. If $|x| = |y|$ or $|x| > |y|$ and $x \in ya\Sigma^*$, then for $k \geq |x| \geq |y|$, $(a^kxb)^i(a^kyb)^j \in D(1)$ and $(a^kyb)^j(a^kxb)^i \in D(1)$ for all $i, j \geq 1$.*

Similarly, in Proposition 5 we show the relationship between some catenations of powers of two θ-unbordered words and the set of all θ-unbordered words.

Proposition 5 *Let θ be a morphic involution on Σ^* and let $a, b \in \Sigma$ such that $a \notin \{\theta(a), b\}$. Let $x, y \in \theta(b)\Sigma^*$ with $x \neq y$. Then for all $k > |x| \geq |y|$, $i, j \geq 1$, $(a^kx\theta(b))^i(a^ky\theta(b))^j \in D_\theta(1)$ and $(a^ky\theta(b))^j(a^kx\theta(b))^i \in D_\theta(1)$.*

Proof To prove the result, we will use Lemma 11 of [19] which states that $\theta(\mathrm{Pref}(u)) \cap \mathrm{Suff}(v) = \emptyset$ and the set of all words in u^+v^+ are θ-unbordered are equivalent statements. Hence we need to show that,

$$\theta(\mathrm{Pref}(a^ky\theta(b))) \cap \mathrm{Suff}(a^kx\theta(b)) = \emptyset \text{ and } \theta(\mathrm{Pref}(a^kx\theta(b))) \cap \mathrm{Suff}(a^ky\theta(b)) = \emptyset.$$

First, let $|x| = |y|$. Then, from Lemma 6 and since $x, y \in \theta(b)\Sigma^*$, we have that, $(a^ky\theta(b))(a^kx\theta(b)) \in D_\theta(1)$ and $(a^kx\theta(b))(a^ky\theta(b)) \in D_\theta(1)$. Therefore, if $|x| = |y|$, then $\theta(\mathrm{Pref}(a^kx\theta(b))) \cap \mathrm{Suff}(a^ky\theta(b)) = \emptyset$ and $\theta(\mathrm{Pref}(a^ky\theta(b))) \cap \mathrm{Suff}(a^kx\theta(b)) = \emptyset$.

Now, let $|x| > |y|$. We will only prove that $\theta(\mathrm{Pref}(a^kx\theta(b))) \cap \mathrm{Suff}(a^ky\theta(b)) = \emptyset$, since the other equality can be proved similarly. Let us assume that $\theta(\mathrm{Pref}(a^kx\theta(b))) \cap \mathrm{Suff}(a^ky\theta(b)) \neq \emptyset$, i.e., there exists $w \in \Sigma^+$ such that $w \in \theta(\mathrm{Pref}(a^kx\theta(b))) \cap \mathrm{Suff}(a^ky\theta(b))$, i.e. $\theta(w) <_d^\theta (a^kx\theta(b))(a^ky\theta(b))$. By Lemma 3, it is enough to consider only the cases when $1 \leq |w| \leq k + |x|$.
Case 1: $|w| < k$. Then $w = \theta(a^n) = y''\theta(b)$ for some $1 \leq n < k$ and $y = y'y''$, for $y' \in \Sigma^+$ and $y'' \in \Sigma^*$ which implies $\theta(a) = \theta(b)$, i.e., $a = b$, a contradiction.
Case 2: $k \leq |w| \leq k + |x|$. Then $w = \theta(a^k)\theta(x') = a^ny\theta(b)$ for some $1 \leq n \leq k$ and $x = x'x''$, $x', x'' \in \Sigma^*$ which implies $\theta(a) = a$, a contradiction.

Since both the cases lead to a contradiction, we have that

$$\theta(\mathrm{Pref}(a^kx\theta(b))) \cap \mathrm{Suff}(a^ky\theta(b)) = \emptyset$$

Similarly, we can prove that $\theta(\mathrm{Pref}(a^ky\theta(b))) \cap \mathrm{Suff}(a^kx\theta(b)) = \emptyset$. Hence,

$$(a^ky\theta(b))^i(a^kx\theta(b))^j, (a^kx\theta(b))^j(a^ky\theta(b))^i \in D_\theta(1).$$

□

We will illustrate Proposition 5 with the following example.

Example 2 Let $\Sigma = \{A, C, G, T\}$ and θ be a morphic involution such that $\theta(A) = T$, $\theta(G) = C$ and viceversa. Let $k = 3$, $i = 2$, $j = 1$ and let $x = TAG, y = TC$. Since $x \neq y$, $x, y \in T\Sigma^*$, $\theta(a) \neq a$ for all $a \in \Sigma$ and $k > i > j$, we have that,

$$(GGGTAGT)^2(GGGTCT) = GGGTAGTGGGTAGTGGGTCT \in D_\theta(1) \text{ and}$$

$$(GGGTCT)(GGGTAGT)^2 = GGGTCTGGGTAGTGGGTAGT \in D_\theta(1).$$

The following lemma is needed for proving the main result of this section.

Lemma 7 *[18] Let θ be a literal (anti)morphism on Σ^* and let $a, b \in \Sigma$ such that $a \neq \theta(b)$. Let $x \neq y$, $x, y \in \Sigma^m$, $m > 0$. Then:*

1. *$a^m x\theta(b) \in D(1)$.*
2. *If $x = \theta(b)x'$, $x' \in \Sigma^*$ and $k > m$, then $(a^k y\theta(b))(a^k x\theta(b)) \in D(1)$.*

Now, we will prove one of the main results of the section which shows that, under certain conditions, the set $D^i_\theta(1)\backslash D(i)$ is disjunctive for all $i \geq 2$, alphabet Σ, and morphic involutions θ.

Theorem 4 *Let $|\Sigma| \geq 3$ and θ be a morphic involution on Σ^* such that $\theta(a) \neq a$ for some $a \in \Sigma$. Then $D^i_\theta(1)\backslash D(i)$ is disjunctive for all $i \geq 2$.*

Proof Since $|\Sigma| \geq 3$ and $\theta(a) \neq a$ for all $a \in \Sigma$ there exists $c \neq a$ such that $\theta(a) = c$ and $\theta(c) = a$. Also, since $|\Sigma| \geq 3$, there exists $b \in \Sigma$ such that $b \neq a$, $b \neq c = \theta(a)$. Let $x, y \in \Sigma^n$, $x \neq y$, $m = n + 1$, $n > 0$. Choose $u = (a^m\theta(b)x\theta(b))^{i-1}a^m\theta(b)$ and $v = \theta(b)$. Since $a \neq b$, by Lemma 6, $a^m\theta(b)x\theta(b) \in D_\theta(1)$. Also, since $a \neq \theta(b)$, $a^m\theta(b)x\theta(b) \in Q$ and by Lemma 7, $a^m\theta(b)x\theta(b) \in D(1)$. Hence by Proposition 2, $\nu_d((a^m\theta(b)x\theta(b))^i) = i$. Thus,

$$uxv = (a^m\theta(b)x\theta(b))^i \in D^i_\theta(1) \cap D(i).$$

On the other hand, by Proposition 4, since $|x| = |y|$, $(a^m\theta(b)x\theta(b))^{i-1}(a^m\theta(b)y\theta(b)) \in D(1)$ and hence

$$uyv = (a^m\theta(b)x\theta(b))^{i-1}(a^m\theta(b)y\theta(b)) \in D^i_\theta(1)\backslash D(i) \text{ for } i \geq 2.$$

Thus, $x \not\equiv y(P_{D^i_\theta(1)\backslash D(i)})$ for every $x \neq y$, $|x| = |y|$ and $i \geq 2$. Hence, by Proposition 1, $D^i_\theta(1)\backslash D(i)$ is disjunctive for all $i \geq 2$. ☐

We know from [30] that $D(1)\backslash\Sigma \subseteq D^2(1)$. Moreover, $D(i)\backslash\Sigma^i \not\subseteq D^{i+1}(1)$ for $i \geq 2$, see [31]. However, for a morphic involution θ, we have that $D_\theta(1)\backslash\Sigma \not\subseteq D^2_\theta(1)$. For example, let $\{a, b\} \in \Sigma$ be such that $\theta(a) = b$ and $\theta(b) = a$. Then $w = abaa \in D_\theta(1)$ but there does not exist any $u, v \in D_\theta(1)$ such that $w = uv$ and $u, v \neq \lambda$. Theorem 5 establishes, under certain conditions, the relationship between $D_\theta(i)$ and $D_\theta(1)$ for $i \geq 2$. The following is a known result, here with a different proof.

Lemma 8 *[15] Let θ be a morphic involution on Σ^* and $u \in \Sigma^+$. Then $u \in D_\theta(1)$ if and only if $\theta(u) \in D_\theta(1)$.*

Proof Let $u \in D_\theta(1)$. Assume that $\theta(u) \notin D_\theta(1)$. Then there exists $v, \alpha, \beta \in \Sigma^+$ such that $\theta(u) = v\alpha = \beta\theta(v)$ which implies $u = \theta(v)\theta(\alpha) = \theta(\beta)v$ which further implies $u \notin D_\theta(1)$, a contradiction. Hence $\theta(u) \in D_\theta(1)$. Similarly, we can prove that if $\theta(u) \in D_\theta(1)$ then $u \in D_\theta(1)$. $\square$

Theorem 5 *Let θ be a morphic involution on Σ^*, where Σ is an alphabet with $|\Sigma| > 2$ that contains letters $a \neq b$ such that $a \neq \theta(b)$ and $\theta(a) \neq a$ for all $a \in \Sigma$. Then the set $D_\theta^{2i}(1)\backslash D_\theta(i+1)$ is disjunctive for all $i \geq 1$.*

Proof Let $x, y \in \Sigma^n$, $x \neq y$, $m = n+1$, $n > 0$. Let $u = a^m\theta(b)$,

$$v = \theta(b)(\theta(a^m\theta(b)x\theta(b))a^m\theta(b)x\theta(b))^{i-1}\theta(a^m\theta(b)x\theta(b)).$$

Since $a \neq b$, by Lemma 6, we have $a^m\theta(b)x\theta(b) \in D_\theta(1)$. Hence by Lemma 8, $\theta(a^m\theta(b)x\theta(b)) \in D_\theta(1)$. Thus, by Proposition 3,

$$uxv = (a^m\theta(b)x\theta(b)\theta(a^m\theta(b)x\theta(b)))^i \in D_\theta(i+1) \cap D_\theta^{2i}(1).$$

Further by Lemma 5,

$$uyv = a^m\theta(b)y\theta(b)(\theta(a^m\theta(b)x\theta(b))a^m\theta(b)x\theta(b))^{i-1}\theta(a^m\theta(b)x\theta(b)) \in D_\theta(1)$$

and hence $uyv \in D_\theta^{2i}(1)\backslash D_\theta(i+1)$ for $i \geq 1$. Therefore, $x \not\equiv y(P_{D_\theta^{2i}(1)\backslash D_\theta(i+1)})$ for every $x, y \in \Sigma^+$, $x \neq y$, $|x| = |y|$ and $i \geq 1$. By Proposition 1, $D_\theta^{2i}(1)\backslash D_\theta(i+1)$ is disjunctive for $i \geq 1$. $\square$

5 Disjunctivity of the set $(D_\theta(2) \cap D(2))\backslash(D_\theta(1) \cap D(1))^k$ for $k = 1, 2$

We have already discussed some relationships between the sets $D_\theta(i)$ and $D(i)$. In particular, the intersection of these two sets is a non-empty set and that, under certain conditions, the sets $D_\theta^i(1)\backslash D(i)$ are disjunctive for all $i \geq 2$. A natural question that arises in this context is what are the relationships between the sets $D_\theta(i) \cap D(i)$ for different values of $i \geq 1$. In this section, we will show that under certain conditions the set $(D_\theta(2) \cap D(2))\backslash(D_\theta(1) \cap D(1))^k$ is disjunctive for $k = 1, 2$ (Theorem 6).

In order to prove the disjunctivity of the set $(D_\theta(2) \cap D(2))\backslash(D_\theta(1) \cap D(1))^k$, $k = 1, 2$, we need to characterize a word or set of words that have exactly two borders and two θ-borders. The following proposition provides such characterization.

Proposition 6 *Let θ be a morphic involution such that $\theta(a) \neq a$ for all $a \in \Sigma$, let $u \in D_\theta(1) \cap D(1)$ and let $u' \in Pref(u) \cup Suff(u) \cup \theta(Suff(u))$ with $u' \neq u$. Then $w = u\theta(u)u'\theta(u)u \in D_\theta(2) \cap D(2)$.*

Proof Let $w = u\theta(u)u'\theta(u)u$. Let us assume that $u' \in \mathrm{Pref}(u)$. Clearly, $\{\lambda, u\} \subseteq L_d(w)$ and $\{\lambda, u\theta(u)\} \subseteq L_d^\theta(u)$. Now, we need to show that $L_d(w) \subseteq \{\lambda, u\}$ and $L_d^\theta(u) \subseteq \{\lambda, u\theta(u)\}$. Since $a \neq \theta(a)$ for all $a \in \Sigma$, $u\theta(u) \notin L_d(w)$ and $u \notin L_d^\theta(u)$. Let us assume that $x \in L_d(w)$ or $y \in L_d^\theta(w)$, i.e. $x <_d w$ or $y <_d^\theta w$ such that $x, y \notin \{\lambda, u, u\theta(u)\}$ and let, $|u| = m, |u'| = n$ where $m > n > 0$. Since $x, y \notin \{u, u\theta(u)\}$, the cases $|x| = |y| = m$ and $|x| = |y| = 2m$ are not possible. Thus we have following 5 cases to consider.

Case 1: If $0 < |x| < m$, then $x \in \mathrm{Pref}(u) \cap \mathrm{Suff}(u)$ which implies $u \notin D(1)$, a contradiction.

If $0 < |y| < m$, then $y \in \mathrm{Pref}(u) \cap \theta(\mathrm{Suff}(u))$ which implies $u \notin D_\theta(1)$, a contradiction.

Case 2: If $m < |x| < 2m$, then for $u = u_1u_2 = u_1'u_2'$ and $|u_1| = |u_2'|$, we will get $x = u\theta(u_1) = \theta(u_2')u = u_1'u_2'\theta(u_1) = \theta(u_2')u_1'u_2'$ where $u_1, u_2, u_1', u_2' \in \Sigma^+$ which implies $\theta(u_1) = u_2'$ which further implies $u \notin D_\theta(1)$, a contradiction.

If $m < |y| < 2m$, then for $u = u_3u_4 = u_3'u_4'$ and $|u_3| = |u_4'|$, we will get $y = u\theta(u_3) = u_4'\theta(u) = u_3'u_4'\theta(u_3) = u_4'\theta(u_3')\theta(u_4')$ where $u_3, u_4, u_3', u_4' \in \Sigma^+$ which implies $\theta(u_3) = \theta(u_4')$, i.e., $u_3 = u_4'$ which further implies $u \notin D(1)$, a contradiction.

Case 3: If $2m < |x| \leq 2m+n$, then $x = u\theta(u)u_1' = u_2'\theta(u)u$ where $u_1' \leq_p u' <_p u$ and $u_2' \leq_s u'$ for $u_1', u_2' \in \Sigma^+$. Since, $|u_1'| \leq |u'| < |u|$, $u_1' <_s u$ which implies $u \notin D(1)$, a contradiction.

If $2m < |y| \leq 2m + n$, then $y = u\theta(u)u_3' = \theta(u_4')u\theta(u)$ where $u_3' \leq_p u' <_p u$ and $u_4' \leq_s u'$ for $u_3', u_4' \in \Sigma^+$. Since $|u_3'| \leq |u'| < |u|$, $u_3' <_s \theta(u)$, i.e., $\theta(u_3') <_s u$ which implies $u \notin D_\theta(1)$, a contradiction.

Case 4: If $2m + n < |x| \leq 3m + n$. Then $x = u\theta(u)u'\theta(u_1) = \theta(u_2)u'\theta(u)u$ where $u_1 \leq_p u$ and $u_2 \leq_s u$ for $u_1, u_2 \in \Sigma^+$. Since, $|u_1| \leq |u|$, $\theta(u_1) \leq_s u$ which implies $u \notin D_\theta(1)$, a contradiction.

If $2m + n < |y| \leq 3m + n$. Then $y = u\theta(u)u'\theta(u_3) = u_4\theta(u')u\theta(u)$ where $u_3 \leq_p u$ and $u_4 \leq_s^\theta u$ for $u_3, u_4 \in \Sigma^+$. Since, $|u_3| \leq |u|$, $\theta(u_3) \leq_s \theta(u)$, i.e., $u_3 \leq_s u$ which implies $u \notin D(1)$, a contradiction.

Case 5: If $3m + n < |x| < 4m + n$. Then $x = u\theta(u)u'\theta(u)u_1 = u_2\theta(u)u'\theta(u)u$ where $u_1 <_p u$, $u_2 <_s u$ for $u_1, u_2 \in \Sigma^+$. Since, $|u_1| < |u|$, $u_1 <_s u$, which implies $u \notin D(1)$, a contradiction.

If $3m + n < |y| < 4m + n$. Then $y = u\theta(u)u'\theta(u)u_3 = \theta(u_4)u\theta(u')u\theta(u)$ where $u_3 <_p u$, $\theta(u_4) <_s u$ for $u_3, u_4 \in \Sigma^+$. Since, $|u_3| < |u|$, $u_3 <_s \theta(u)$, i.e., $\theta(u_3) <_s u$ which implies $u \notin D_\theta(1)$, a contradiction.

If we assume that $u' \in \mathrm{Suff}(u)$ or $u' \in \theta(\mathrm{Suff}(u))$, we will reach a similar contradiction.

Since all the cases lead to a contradiction, $w \in D_\theta(2) \cap D(2)$. □

We illustrate Proposition 6 with the following example.

Example 3 Let $\Sigma = \{A, C, G, T\}$ and θ be a morphic involution such that $\theta(A) = T$, $\theta(G) = C$ and vice versa. Let $u = GTA$ and $u' = GT$. Then for $w = u\theta(u)u'\theta(u)u = GTACATGTCATGTA$, $L_d^\theta(w) = \{\lambda, GTACAT\}$ and $L_d(w) = \{\lambda, GTA\}$. Hence $w \in D_\theta(2) \cap D(2)$.

Similarly, we need to prove that the words of the form mentioned in Proposition 6 cannot be decomposed as a catenation of less than three words which are unbordered as well as θ-unbordered.

Proposition 7 *Let θ be a morphic involution such that $\theta(a) \neq a$ for all $a \in \Sigma$ and let $u \in D_\theta(1) \cap D(1)$, $u' \in Pref(u) \cup Suff(u) \cup \theta(Suff(u))$ such that $u' \neq u$, then $u\theta(u)u'\theta(u)u \notin (D_\theta(1) \cap D(1))^n$ for $1 \leq n \leq 2$.*

Proof Let us assume that $w = u\theta(u)u'\theta(u)u \in (D_\theta(1) \cap D(1))^n$ for $1 \leq n \leq 2$. Let us assume that $u' \in \mathrm{Pref}(u)$. The case $n = 1$ is not possible since by

Proposition 6, $w \in D_\theta(2) \cap D(2)$. Hence, let $n = 2$, i.e., $u\theta(u)u'\theta(u)u = v_1v_2$ where $v_1, v_2 \in D_\theta(1) \cap D(1)$. Then we have following cases to consider.

Case 1: $v_1 = u$, $v_2 = \theta(u)u'\theta(u)u$. Then $v_2 \in D_\theta(2)$, a contradiction.

Case 2: $v_1 = u\theta(u)$, $v_2 = u'\theta(u)u$. Then, $v_1 \in D_\theta(2)$, a contradiction.

Case 3: $v_1 = u\theta(u)u'$, $v_2 = \theta(u)u$. Then $v_2 \in D_\theta(2)$, a contradiction.

Case 4: $v_1 = u\theta(u)u'\theta(u)$, $v_2 = u$. Then $v_1 \in D_\theta(2)$, a contradiction.

Case 5: $v_1 = u_1$, $v_2 = u_2\theta(u)u'\theta(u)u$ where $u = u_1u_2$ and $u_1, u_2 \in \Sigma^+$. This implies $v_2 = u_2\theta(u)u'\theta(u)u_1u_2 \in D(2)$, a contradiction.

Case 6: $v_1 = u\theta(x_1)$, $v_2 = \theta(x_2)u'\theta(u)u$ where $u = x_1x_2$ and $x_1, x_2 \in \Sigma^+$. This implies, $v_2 = \theta(x_2)u'\theta(u)x_1x_2 \in D_\theta(2)$, a contradiction.

Case 7: $v_1 = u\theta(u)u'_1$, $v_2 = u'_2\theta(u)u$ where $u' = u'_1u'_2$ and $u'_1, u'_2 \in \Sigma^+$. Also, since $u' <_p u$, $u = u'u'_3 = u'_1u'_2u'_3$ where $u'_3 \in \Sigma^+$. This implies, $v_1 = u'_1u'_2u'_3\theta(u)u'_1, \in D(2)$, a contradiction.

Case 8: $v_1 = u\theta(u)u'\theta(u_1)$, $v_2 = \theta(u_2)u$ where $u = u_1u_2$ and $u_1, u_2 \in \Sigma^+$ which implies, $v_1 = u_1u_2\theta(u)u'\theta(u_1) \in D_\theta(2)$, a contradiction.

Case 9: $v_1 = u\theta(u)u'\theta(u)u_1$, $v_2 = u_2$ where $u = u_1u_2$ and $u_1, u_2 \in \Sigma^+$, which implies, $v_1 = u_1u_2\theta(u)u'\theta(u)u_1 \in D(2)$, a contradiction.

If we assume that $u' \in \mathrm{Suff}(u)$ or $u' \in \theta(\mathrm{Suff}(u))$, we will reach a similar contradiction.

Since all cases led to contradictions, $w \notin (D_\theta(1) \cap D(1))^n$ for $1 \le n \le 2$. □

As a consequence of Proposition 6 and 7, we have the following result.

Corollary 2 *Let θ be a morphic involution such that $\theta(a) \neq a$ for all $a \in \Sigma$ and let $u \in D_\theta(1) \cap D(1)$, $u' \in$ Pref$(u) \cup$ Suff$(u) \cup \theta($Suff$(u))$ such that $u' \neq u$. Then $u\theta(u)u'\theta(u)u \in (D_\theta(2) \cap D(2)) \backslash (D_\theta(1) \cap D(1))^n$ for $1 \le n \le 2$.*

The following result is needed for the proof of Lemma 10.

Lemma 9 *[18] Let θ be a morphic involution on Σ^*, where Σ is an alphabet with $|\Sigma| \ge 3$ that contains letters $a \neq b$ such that $a \notin \{\theta(b), \theta(a)\}$. Let $x \neq y$, $x, y \in \Sigma^m$, $m > 0$. Then:*

1. *$a^m x\theta(b) \in D_\theta(1) \cap D(1)$.*
2. *If $x = \theta(b)x'$, $x' \in \Sigma^*$ and $k \ge m$, then $(a^k y\theta(b))(a^k x\theta(b)) \in D_\theta(1) \cap D(1)$.*

In the following lemma we prove that, for a morphic involution θ, certain words of the form $u\theta(u)u'\theta(u)v$, where $u' <_p u$ and $u \neq v$, are unbordered as well as θ-unbordered.

Lemma 10 *Let $|\Sigma| \ge 3$, θ be a morphic involution with the property that there exists $a \in \Sigma$ such that $a \notin \{\theta(a), b, \theta(b)\}$. Then for $u, v \in \Sigma^n$ such that $u \neq v$, $n > 0$ and $m = n + 2$,*

$$a^m\theta(b)u\theta(b)\theta(a^m\theta(b)u\theta(b))a^m\theta(a^m\theta(b)u\theta(b))a^m\theta(b)v\theta(b) \in D_\theta(1) \cap D(1).$$

Proof Let us assume that

$$w = a^m\theta(b)u\theta(b)\theta(a^m\theta(b)u\theta(b))a^m\theta(a^m\theta(b)u\theta(b))a^m\theta(b)v\theta(b) \notin D_\theta(1) \cap D(1),$$

i.e., there exists $w_1, w_2 \in \Sigma^+$ such that $w_1 <_d w$ or $w_2 <_d^\theta w$. By Lemma 2 and Lemma 3, it is enough consider the case $|w_1| < 5m$ or $|w_2| < 5m$. Further by

Lemma 9, taking $y = \theta(b)u$ and $x = \theta(b)v$ we know that $(a^m\theta(b)u\theta(b))(a^m\theta(b)v\theta(b)) \in D_\theta(1) \cap D(1)$, hence none of the prefixes of $a^m\theta(b)u\theta(b)$ can be a border or a θ-border of w and hence the cases $1 \leq |w_1| \leq 2m$ or $1 \leq |w_2| \leq 2m$ are not possible. So, we only need to consider the cases when $2m < |w_1| < 5m$ or $2m < |w_2| < 5m$.

Case 1: $2m < |w_1| < 3m$ or $2m < |w_2| < 3m$. Then,

$$w_1 = a^m\theta(b)u\theta(b)\theta(a^k) = \theta(u_2)ba^m\theta(b)v\theta(b) \text{ or}$$

$$w_2 = a^m\theta(b)u\theta(b)\theta(a^{k'}) = u_2'\theta(b)\theta(a^m)b\theta(v)b$$

where $1 \leq k, k' < m$, $u = u_1u_2 = u_1'u_2'$, $u_1, u_1' \in \Sigma^+$ and $u_2, u_2' \in \Sigma^*$ which implies $a = b$ or $a = \theta(b)$, a contradiction.

Case 2: $|w_1| = 3m$ or $|w_2| = 3m$. Then,

$$w_1 = a^m\theta(b)u\theta(b)\theta(a^m) = b\theta(u)ba^m\theta(b)v\theta(b) \text{ or}$$

$$w_2 = a^m\theta(b)u\theta(b)\theta(a^m) = \theta(b)u\theta(b)\theta(a^m)b\theta(v)b$$

which implies $a = b$ or $a = \theta(b)$, a contradiction.

Case 3: $|w_1| = 3m + 1$ or $|w_2| = 3m + 1$. Then,

$$w_1 = a^m\theta(b)u\theta(b)\theta(a^m)b = \theta(a)b\theta(u)ba^m\theta(b)v\theta(b) \text{ or}$$

$$w_2 = a^m\theta(b)u\theta(b)\theta(a^m)b = a\theta(b)u\theta(b)\theta(a^m)b\theta(v)b$$

which implies $a = \theta(a)$ (and $a = b$) or $a = \theta(b)$, a contradiction.

Case 4: $3m + 1 < |w_1| \leq 4m - 1$ or $3m + 1 < |w_2| \leq 4m - 1$. Then,

$$w_1 = a^m\theta(b)u\theta(b)\theta(a^m)b\theta(u_1) = \theta(a^k)b\theta(u)ba^m\theta(b)v\theta(b) \text{ or}$$

$$w_2 = a^m\theta(b)u\theta(b)\theta(a^m)b\theta(u_1') = a^{k'}\theta(b)u\theta(b)\theta(a^m)b\theta(v)b$$

where $2 \leq k, k' \leq n + 1$, $u = u_1u_2 = u_1'u_2'$, $u_1, u_1' \in \Sigma^+$ and $u_2, u_2' \in \Sigma^*$ which implies $a = b$ (and $\theta(a) = a$) or $a = \theta(b)$, a contradiction.

Case 5: $|w_1| = 4m$ or $|w_2| = 4m$. Then,

$$w_1 = a^m\theta(b)u\theta(b)\theta(a^m)b\theta(u)b = \theta(a^m)b\theta(u)ba^m\theta(b)v\theta(b) \text{ or}$$

$$w_2 = a^m\theta(b)u\theta(b)\theta(a^m)b\theta(u)b = a^m\theta(b)u\theta(b)\theta(a^m)b\theta(v)b$$

which implies $a = \theta(a)$ or $\theta(u) = \theta(v)$, i.e., $u = v$, a contradiction.

Case 6: $4m < |w_1| < 5m$ or $4m < |w_2| < 5m$. Then,

$$w_1 = a^m\theta(b)u\theta(b)\theta(a^m)b\theta(u)ba^k = a^{k'}\theta(a^m)b\theta(u)ba^m\theta(b)v\theta(b) \text{ or}$$

$$w_2 = a^m\theta(b)u\theta(b)\theta(a^m)b\theta(u)ba^{k_1} = \theta(a^{k_2})a^m\theta(b)u\theta(b)\theta(a^m)b\theta(v)b$$

where $1 < k, k', k_1, k_2 < m$ which implies $a = \theta(b)$ (and $a = \theta(a)$) or $a = b$ (and $a = \theta(a)$), a contradiction.

Since all the cases lead to a contradiction $w \in D_\theta(1) \cap D(1)$. □

The following theorem uses Lemma 10, along with certain conditions on the alphabet Σ, to show the disjunctivity of the languages $(D_\theta(2) \cap D(2)) \backslash (D_\theta(1) \cap D(1))^k$ for $k = 1, 2$.

Theorem 6 *Let $|\Sigma| \geq 3$ and θ be a morphic involution such that $\theta(a) \neq a$ for all $a \in \Sigma$. Then $(D_\theta(2) \cap D(2)) \backslash (D_\theta(1) \cap D(1))^k$ is disjunctive for $k = 1, 2$.*

Proof Let $\{a, b\} \in \Sigma$ such that $a \notin \{b, \theta(b)\}$. Let $u, v \in \Sigma^n$ for $n > 0$ be such that $u \neq v$. Let $m = |bub| = n + 2$. Now let,

$$x = a^m\theta(b)u\theta(b)\theta(a^m\theta(b)u\theta(b))a^m\theta(a^m\theta(b)u\theta(b))a^m\theta(b)$$

and $y = \theta(b)$. Then

$$xuy = a^m\theta(b)u\theta(b)\theta(a^m\theta(b)u\theta(b))a^m\theta(a^m\theta(b)u\theta(b))a^m\theta(b)u\theta(b) \text{ and}$$

$$xvy = a^m\theta(b)u\theta(b)\theta(a^m\theta(b)u\theta(b))a^m\theta(a^m\theta(b)u\theta(b))a^m\theta(b)v\theta(b).$$

By Corollary 2, $xuy \in (D_\theta(2) \cap D(2)) \backslash (D_\theta(1) \cap D(1))^k$ for $k = 1, 2$. Now, by Lemma 10, we know that $xvy \in D_\theta(1) \cap D(1)$. Therefore, $u \not\equiv v(P_L)$ for every $u, v \in \Sigma^+$, $u \neq v$, $|u| = |v|$ and $L = (D_\theta(2) \cap D(2)) \backslash (D_\theta(1) \cap D(1))^k$ for $k = 1, 2$ is, by Proposition 1, disjunctive. □

We conclude this section with some observations on the disjunctivity of some other languages related to $D_\theta(i)$, $i \geq 1$. Let us recall the definition of a singular language from [26]. For any language $L \subseteq \Sigma^+$, [26] defines,

$$l(L) = \{g \in L | gx \notin L \text{ for all } x \in \Sigma^+ \text{ and } g = yz, z \in \Sigma^+, \text{ implies } y \notin L\}.$$

Each element of $l(L)$ is called a *singular word in* L and L is said to be a *singular language* if $l(L) \neq \emptyset$.

Theorem 7 *[26] Let L' be a disjunctive language and let L be a singular language. Then LL' is a disjunctive language.*

Corollary 3 *If Σ is such that $|\Sigma| > 2$ and $a \neq \theta(a)$ for all $a \in \Sigma$, θ is a morphic involution on Σ^*, and L is a singular language over Σ, then the following hold:*

1. *If there exist $\{a, b\} \in \Sigma$ such that $a \notin \{b, \theta(b)\}$ then $LD_\theta(i)$ is disjunctive for all $i \geq 1$.*
2. *If there exist $\{a, b\} \in \Sigma$ such that $a \notin \{b, \theta(b)\}$ then $L(D_\theta(i) \cap Q_\theta^{2i-2})$ is disjunctive for all $i \geq 2$.*
3. *The language $L(D_\theta^i(1) \backslash D(i))$ is disjunctive for all $i \geq 2$.*
4. *If there exist $\{a, b\} \in \Sigma$ such that $a \notin \{b, \theta(b)\}$ then $L(D_\theta^{2i}(1) \backslash D_\theta(i + 1))$ is disjunctive for all $i \geq 1$.*
5. *The language $L((D_\theta(2) \cap D(2)) \backslash (D_\theta(1) \cap D(1))^k)$ is disjunctive for $k = 1, 2$.*

Proof 1. By Theorems 3 and 7, $LD_\theta(i)$ is disjunctive for $i \geq 1$.
2. By Corollary 1 and Theorem 7, $L(D_\theta(i) \cap Q_\theta^{2i-2})$ is disjunctive for all $i \geq 2$.
3. By Theorems 4 and 7, $L(D_\theta^i \backslash D(i))$ is disjunctive for all $i \geq 2$.
4. By Theorems 5 and 7, $L(D_\theta^{2i}(1) \backslash D_\theta(i + 1))$ is disjunctive for all $i \geq 1$.
5. By Theorems 6 and 7, $L((D_\theta(2) \cap D(2)) \backslash (D_\theta(1) \cap D(1))^k)$ is disjunctive for $k = 1, 2$.

□

6 Further remarks on $D_\theta(i)$ and related languages

As seen in Section 4, for a word $u \in D_\theta(1)$ there might not exist a decomposition $u = u_1u_2$ such that $u_1, u_2 \in D_\theta(1)$. If, however, such a decomposition exists for a non-empty word, then that word is said to be $D_\theta(1)$-*concatenate*. The word u is said to be *completely* $D_\theta(1)$-*concatenate*, if $u = xy$ for $x, y \in \Sigma^+$, imply that $x, y \in D_\theta(1)$. These notions generalize concepts related to $D(1)$-*concatenate* words, defined in [14].

Example 4 Let $\Sigma = \{a, b\}$, and θ be (anti)morphic involution such that $\theta(a) = b$ and vice versa. Then $u = ab$ is $D_\theta(1)$-concatenate. Also, $v = a^i$, $i \geq 1$ is completely $D_\theta(1)$-concatenate, but $w = aba = (a\theta(a))(a)$ is not $D_\theta(1)$-concatenate and hence not completely $D_\theta(1)$-concatenate.

The following proposition shows that the set of all completely $D_\theta(1)$-concatenate words is regular for an (anti)morphic involution θ.

Proposition 8 *Let Σ be an alphabet, L be the set of all completely $D_\theta(1)$-concatenate words over Σ, and let θ be an (anti)morphic involution. Then L is regular.*

Proof Let $\Sigma = \{a_1, a_2, \ldots, a_n\}$. Let $u \in L$ be such that $u = a_iwa_j$, $1 \leq i, j \leq n$, $w \in \Sigma^*$. If w does not contain $\theta(a_i)$ and $\theta(a_j)$, then for $w = w'w''$, $w', w'' \in \Sigma^*$, $a_iw', w''a_j \in D_\theta(1)$. On the other hand, if w contains $\theta(a_i)$ or $\theta(a_j)$, then for some $w', w'' \in \Sigma^*$, $w = w'\theta(a_i)w''$ or $w = w'\theta(a_j)w''$. Thus we have, $u = (a_iw'\theta(a_i))w''a_j$ or $a_iw'(\theta(a_j)w''a_j)$, which contradicts to the fact that u is completely $D_\theta(1)$-concatenate. Thus,

$$L = \bigcup_{i=1,j=1}^{n} a_i(\Sigma \backslash \{\theta(a_i), \theta(a_j)\})^* a_j.$$

Since $a_i(\Sigma \backslash \{\theta(a_i), \theta(a_j)\})^* a_j$ is regular, L is regular. □

The catenation of θ-unbordered words is not necessarily θ-unbordered. Additional conditions, such as the one below, are needed to guarantee that the catenation of θ-unbordered words is θ-unbordered.

Proposition 9 *[19] Let θ be either a morphic or an antimorphic involution and let $u, v \in \Sigma^+$ be θ-unbordered. Then uv is θ-unbordered iff $\theta(Pref(u)) \cap Suff(v)) = \emptyset$.*

Based on above proposition and the notion of non-overlapped languages defined in [31], we now introduce a new class of languages, called θ-non-overlapped languages. θ-non-overlapped languages are a special class of θ-unbordered words, whose additional properties imply that the catenation between any two words in the language remains θ-unbordered.

A pair of words $u, v \in \Sigma^+$, $u \neq v$, is said to be θ-*non-overlapped* iff $\theta(\mathrm{Pref}(u)) \cap \mathrm{Suff}(v) = \emptyset$ and $\theta(\mathrm{Pref}(v)) \cap \mathrm{Suff}(u) = \emptyset$. A language $L \subseteq \Sigma^+$ is said to be θ-*non-overlapped* if $L \subseteq D_\theta(1)$ and $u, v \in L$, $\theta(u) \neq v$, implies u and v are θ-non-overlapped.

For a language L, let us denote $L_\theta^{(2)} = \{u\theta(u) | u \in L\}$. The following results describe some properties of θ-non-overlapped languages.

Lemma 11 *Let θ be morphic involution and L be θ-non-overlapped. Then $\theta(L)$ is also θ-non-overlapped.*

The following proposition shows the necessary and sufficient condition for a language to be θ-non-overlapped.

Proposition 10 *Let $L \subseteq \Sigma^+$ and θ be a morphic involution. Then L is θ-non-overlapped language if and only if $L \subseteq D_\theta(1)$ and $L^2 \backslash L_\theta^{(2)} \subseteq D_\theta(1)$.*

Proof Let $L \subseteq \Sigma^+$. Assume that L is a θ-non-overlapped language. Then $L \subseteq D_\theta(1)$. Now, let, $u, v \in L$ such that $v \neq \theta(u)$, i.e. $uv \in L^2 \backslash L_\theta^{(2)}$. Suppose $uv \notin D_\theta(1)$, then there exists $w \in \Sigma^+$ such that $w <_d^\theta uv$. If $|w| > |u|$, then there exists $w' \in \Sigma^+$ such that $w = uw'$ and $uv = uw'\alpha = \beta\theta(u)\theta(w')$ for $\alpha, \beta \in \Sigma^+$. Thus $w' <_d^\theta v$ and $L \not\subseteq D_\theta(1)$, a contradiction. We will reach a similar contradiction if we assume that $|w| > |v|$. If $|w| \leq |u|$ and $|w| \leq |v|$, then $\theta(w) \in \theta(\text{Pref}(u)) \cap \text{Suff}(v)$, a contradiction. Hence $uv \in D_\theta(1)$ and $L^2 \backslash L_\theta^{(2)} \subseteq D_\theta(1)$.

Conversely, assume that $L \subseteq D_\theta(1)$ and $L^2 \backslash L_\theta^{(2)} \subseteq D_\theta(1)$. Consider $u, v \in L$ such that $v \neq \theta(u)$. Then, clearly $uv \in L^2 \backslash L_\theta^{(2)}$. Suppose u, v are not θ-non-overlapped, then $\theta(\text{Pref}(u)) \cap \text{Suff}(v) \neq \emptyset$ or $\theta(\text{Pref}(v)) \cap \text{Suff}(u) \neq \emptyset$. Let $w \in \theta(\text{Pref}(u)) \cap \text{Suff}(v)$ which implies $\theta(w) <_d^\theta uv$. Thus $uv \notin D_\theta(1)$, which is a contradiction to the assumption that $L^2 \backslash L_\theta^{(2)} \subseteq D_\theta(1)$. We will reach a similar contradiction if we assume that $w' \in \theta(\text{Pref}(v)) \cap \text{Suff}(u)$. Thus, $\theta(\text{Pref}(u)) \cap \text{Suff}(v) = \emptyset$ and $\theta(\text{Pref}(v)) \cap \text{Suff}(u) = \emptyset$ for every $u, v \in L$, $\theta(u) \neq v$ and $L \subseteq D_\theta(1)$, i.e. L is θ-non-overlapped. □

We will illustrate Proposition 10 with the following example.

Example 5 Let $\Sigma = \{A, C, G, T\}$ and θ be a morphic involution such that $\theta(A) = T$, $\theta(G) = C$ and viceversa. Let $L = \{AG, GACG\}$, which is a θ-non-overlapped language. Then $L \subseteq D_\theta(1)$ and

$$L^2 \backslash L_\theta^{(2)} = \{AGAG, AGGAGC, GAGCGAGC, GAGCAG\} \in D_\theta(1).$$

Proposition 11 *Let $L \subseteq \Sigma^+$ be a θ-non-overlapped language and let $w \in L^m$ for some $m \geq 1$. If there exists $u \in \Sigma^+$ such that $u \leq_d^\theta w$, then $u \in L^i(\theta(L))^j$ for some $1 \leq i, j, \leq m$.*

Proof Let $w \in L^m$, i.e., $w = w_1 w_2 \ldots w_m$ for some $w_1, w_2, \ldots w_m \in L$. Let $u \in \Sigma^+$ be such that $u \leq_d^\theta w$. Then there exist $1 \leq l \leq m$ such that $u = w_1 \ldots w_{l-1} u_1$, where $u_1 \in \text{Pref}(w_l)$. Thus $u_1 \leq_d^\theta w_l \ldots w_m$. Similarly, there exist $l \leq k \leq m$ such that $\theta(u_1) = u_2 w_{k+1} \ldots w_m$ where $u_2 \in \text{Suff}(w_k)$ which implies $u_1 = \theta(u_2)\theta(w_{k+1}) \ldots \theta(w_m)$. Now, since $\theta(u_2) \leq_p u_1 \leq_p w_l$ and $u_2 \leq_s w_k$, we will get $\theta(u_2) \leq_d^\theta w_l w_k$. Since by Proposition 10, $L^2 \backslash L_\theta^{(2)} \subseteq D_\theta(1)$, $w_k = \theta(w_l)$. Also, since $L, \theta(L) \subseteq D_\theta(1)$, $w_k = \theta(w_l) = u_2$. Thus,

$$\begin{aligned} u &= w_1 \ldots w_{l-1} u_1 \\ &= w_1 \ldots w_{l-1} \theta(u_2) \theta(w_{k+1}) \ldots \theta(w_m) \\ &= w_1 \ldots w_{l-1} w_l \theta(w_{k+1}) \ldots \theta(w_m) \in L^l(\theta(L))^{m-k-1}. \end{aligned}$$

□

For a word $u \in \Sigma^+$,

$$\mathrm{IN}(u) = \{v \in \Sigma^+ | u = xvy \text{ for some } x, y \in \Sigma^*\}.$$

The following result shows the relationship between the length of an infix of a θ-periodic word and the number of borders as well as θ-borders of such infix, for morphic involutions θ.

Proposition 12 *Let $u, v \in \Sigma^+$ and θ be a morphic involution. If $v \in D_\theta(i_1) \cap D(i_2)$ with $i = i_1 + i_2$ for $i_1, i_2 \geq 1$ and $v \in IN(u_1 u_2 \ldots u_m)$ where $u_k = u$ if k is odd and $u_k = \theta(u)$ if k is even for $1 \leq k \leq m$ and $m \geq i$, then $|v| \leq |u^i|$.*

Proof Let us assume that $|v| > |u^i|$. Then there is an integer $i \leq j < m$ such that $|u^j| < |v| \leq |u^{j+1}|$. Hence, v is of the form $v = (v_1 v_2 \ldots v_j)v'$ where $|v_l| = |u|$, $1 \leq l \leq j$, and $v' \in \Sigma^+$, $|v'| \leq |u|$. Since, v is an infix of a word of the form $u_1 u_2 \ldots u_m$ where $u_k = u$ if k is odd and $u_k = \theta(u)$ if k is even for $1 \leq k \leq m$, there exists a word $w \in \Sigma^+$, $|w| = |u|$, such that $v_l = w$ if l is odd and $v_l = \theta(w)$ if l is even, $1 \leq l \leq j$. Furthermore, if j is odd, then $v' \leq_p \theta(w)$ and if j is even, then $v' \leq_p w$. We have the following two cases to consider.

Case 1: j is odd. Then $v = (w\theta(w)w\theta(w)\ldots w)\theta(w')$ for $w = w'\alpha$ where $w' \in \Sigma^+$ and $\alpha \in \Sigma^*$. This implies, $v = (w'\alpha\theta(w')\theta(\alpha)\ldots w'\alpha)\theta(w')$.

Thus, $\lambda, w', (v_1 v_2)w', \ldots, (v_1 v_2 \ldots v_{j-1})w' \in L_d^\theta(v)$ which implies $\nu_d^\theta(v) = \frac{j+3}{2}$.

Also, $\lambda, v_1\theta(w'), (v_1 v_2 v_3)\theta(w'), \ldots, (v_1 v_2 \ldots v_{j-2})\theta(w') \in L_d(v)$. This implies $\nu_d(v) = \frac{j+1}{2}$.

Hence, $\nu_d^\theta(v) + \nu_d(v) = \frac{j+3}{2} + \frac{j+1}{2} = j + 2 \geq i + 2 > i$, which is a contradiction.

Case 2: j is even. Then $v = (w\theta(w)w\theta(w)\ldots\theta(w))w''$ for $w = w''\beta$ where $w'' \in \Sigma^+$ and $\beta \in \Sigma^*$. This implies, $v = (w''\beta\theta(w'')\theta(\beta)\ldots\theta(w'')\theta(\beta))w''$.

Thus, $\lambda, w'', (v_1 v_2)w'', \ldots, (v_1 v_2 \ldots v_{j-2})w'' \in L_d(v)$ which implies $\nu_d(v) = \frac{j+2}{2}$.

Also, $\lambda, v_1\theta(w''), (v_1 v_2 v_3)\theta(w''), \ldots, (v_1 v_2 \ldots v_{j-1})\theta(w'') \in L_d\theta(v)$. This implies $\nu_d(v) = \frac{j+2}{2}$.

Hence, $\nu_d^\theta(v) + \nu_d(v) = \frac{j+2}{2} + \frac{j+2}{2} = j + 2 \geq i + 2 > i$, which is a contradiction.

Since both the cases lead to a contradiction, $|v| \leq |u^i|$. □

We conclude with a preview of possible generalizations of this research to cases where $\theta^3 = I$ over Σ or, more generally, where $\theta^n = I$ over Σ. In [19] it was shown that, for a morphic involution θ, the set of all θ-bordered words over Σ is not context-free. The following results shows that this holds also for the case of a morphism θ with the property that $\theta(a) \neq a$ for all $a \in \Sigma$ and θ^3 equals the identity on Σ with $|\Sigma| \geq 3$.

Proposition 13 *If $|\Sigma| \geq 3$, θ is a morphism such that $\theta^3 = I$ on Σ and $\theta(a) \neq a$ for all $a \in \Sigma$, then the set of all θ-bordered words over Σ is not context-free.*

Proof Let $a \in \Sigma$. Now, since $a \neq \theta(a)$ there exists $c \in \Sigma$ such that $\theta(a) = c$. By the same argument there exists $b \in \Sigma$ such that $\theta(c) = b$. Since, $\theta^3 = I$, $\theta(b) = a$.

Assume that L is context-free. Let n be the constant defined by pumping lemma for context-free languages. Let $w_1 = c^{n+1}a^{n+1}b^{n+1}c^{n+1}$ which is clearly a θ-bordered word. By the pumping lemma, there is a decomposition $w_1 = \alpha x v y \beta$ such that $|xvy| \leq n$, $|xy| \geq 1$ and for all $i \geq 0$, $w_i = \alpha x^i v y^i \beta \in L$. Since w_i begins with c for any $i \geq 0$, every θ-border z of w_i has the property $z = cu$ for some $u \in \Sigma^+$.

Case 1: xvy is a subword of $c^{n+1}a^{n+1}$ of w_1. In this case, since w_i has the suffix c^{n+1}, $\theta(z) \in b\Sigma^* c^{n+1}$. ($\theta(z)$ cannot begin with c or a because in those cases z would begin with a or b respectively, which is not possible.) Hence, $z \in c\Sigma^* a^{n+1}$. If neither x nor y contains any as which means xvy is a subword of c^{n+1} of w_1, we get $w_i = c^m a^{n+1} b^{n+1} c^{n+1}$, for $i \geq 2$ and $m > n+1$. But then, $z = c^m a^{n+1}$ which further imply that $\theta(z_i) = b^m c^{n+1}$, which is a contradiction since w_i does not contain m consecutive bs. Hence, either x or y must include at least one letter a. But this would imply that w_0 has at most n letters a which is a contradiction since it has $z = cua^{n+1}$ for some $u \in \Sigma^*$ as its θ-border.

Case 2: xvy is a subword of $a^{n+1}b^{n+1}$ of w_1. In this case, since w_i has the suffix c^{n+1}, $\theta(z) \in b\Sigma^* c^{n+1}$. Hence, $z \in c\Sigma^* a^{n+1}$. If neither x nor y contains any bs which means xvy is a subword of a^{n+1} of w_1, we get $w_0 = c^{n+1}a^k b^{n+1} c^{n+1}$ for $k \leq n$, which means that w_0 has at most n letters a which contradicts the fact that w_0 has $z_0 = cua^{n+1}$ for some $u \in \Sigma^*$ as its θ-border. Hence, either x or y must include at least one letter b. But then, $w_0 = c^{n+1}a^l b^k c^{n+1} \notin L$ for $k < n+1$ and $l \leq n+1$ since $k < n+1$, $c^{n+1}a^l$ cannot be a θ-border of w_0. Hence we have reached a contradiction

Case 3: xvy is a subword of $b^{n+1}c^{n+1}$ of w_1. In this case, since w_i has prefix c^{n+1}, $z \in c^{n+1}\Sigma^* a$. (z cannot end with c or b because in those cases $\theta(z)$ would end with b or a which is not possible.) Hence, $\theta(z) \in b^{n+1}\Sigma^* c$. If neither x nor y contains any cs which means xvy is a subword of b^{n+1} of w_1, we get $w_0 = c^{n+1}a^{n+1}b^{k'}c^{n+1}$ for $k' \leq n$, which means $w_0 \notin L$ which is a contradiction, because, $c^{n+1}a^{n+1}$ cannot be a θ-border of w_i due to the fact that $k' \leq n$. Hence, either x or y must include at least one letter c. But then, $w_i = c^{n+1}a^{n+1}b^j c^{j'} \notin L$ for $j \geq n+1, j' > n+1$ and $i \geq 2$ since $c^{n+1}a^{n+1}$ cannot be a θ-border of w_i because $j' > n+1$. Hence, we have reached a contradiction.

Lastly, since $|xvy| \leq n$, we have that xvy can also not be a subword of $c^{n+1}a^{n+1}b^{n+1}$ or $a^{n+1}b^{n+1}c^{n+1}$.

Since all the cases lead to a contradiction, our assumption was incorrect and hence L is not context-free. □

In general, the set of all θ-bordered words, B_θ, is not context-free for any morphism θ such that there exists $n \geq 2$ with $\theta^n(a) = a$ for all $a \in \Sigma$. The idea of the proof, [24], is to consider such a morphism and a letter $a \in \Sigma$ such that $\theta^n(a) = a$ for $n > 1$ and $\theta^i(a) \neq a$ for all $0 < i < n$. Now, consider the set $S = B_\theta \cap a^+\theta(a)^+\theta^2(a)^+ \ldots \theta^{n-1}(a)^+a^+$. If $w \in S$, then $w = a^{i_0}(\theta(a))^{i_1}(\theta^2(a))^{i_2} \ldots (\theta^{n-1}(a))^{i_{n-1}}(\theta^n(a))^{i_n}$ where $i_m \geq 1$ for $1 \leq m \leq n$. Let $v \in \Sigma^+$ be such that $v <_d^\theta w$. Thus,

$$\theta(v) = (\theta(a))^j(\theta^2(a))^{i_2} \ldots (\theta^{n-1}(a))^{i_{n-1}}(\theta^n(a))^{i_n}$$

for $j \leq i_1$. Also, $v = a^{i_0}(\theta(a))^{i_1}(\theta^2(a))^{i_2} \ldots (\theta^{n-1}(a))^k$ for $k \leq i_{n-1}$. This implies

$$\theta(v) = (\theta(a))^{i_0}(\theta^2(a))^{i_1} \ldots (\theta^{n-1}(a))^{i_{n-2}}(\theta^n(a))^k.$$

Thus, the comparison of expressions for $\theta(v)$ yields, $i_0 = j \leq i_1$, $i_1 = i_2 = i_3 = \ldots = i_{n-2} = i_{n-1}$ and $i_n = k \leq i_{n-1}$. Hence,

$$S = \{a^{i_0}(\theta(a))^l(\theta^2(a))^l \ldots (\theta^{n-1}(a))^l a^{i_n} | i_0, i_n \leq l\}$$

which is clearly not a context-free language. Thus, if we consider any word from the set S, it will clearly be a θ-bordered word and hence the set B_θ is not context-free.

7 Conclusions

This paper continues the exploration of properties of θ-bordered (pseudo-bordered) words and θ-unbordered words for the case where θ is a morphic involution. We prove, under certain conditions, the disjunctivity of the language of words with exactly i θ-borders, for all $i \geq 1$, and also that the set $D_\theta^i(1)\backslash D(i)$ of the language of words which consist of catenations of i θ-unbordered words, but which do not have exactly i borders, is disjunctive for all $i \geq 2$. Further directions of research include generalizations of these and similar results for morphism or antimorphisms θ with the property that θ^n equals the identity function on Σ for an arbitrary $n \geq 3$.

Acknowledgements We thank Steffen Kopecki for valuable suggestions, in particular those that lead to shortening the proof of Proposition 12, and Sepinoud Azimi for discussions.

References

1. Blondin Massé, A., Gaboury, S., Hallé, S.: Pseudoperiodic words. In: H.C. Yen, O. Ibarra (eds.) Developments in Language Theory, *Lecture Notes in Computer Science*, vol. 7410, pp. 308–319. Springer Berlin Heidelberg (2012)
2. Carpi, A., de Luca, A.: Periodic-like words, periodicity, and boxes. Acta Informatica **37**(8), 597–618 (2001)
3. Cho, D.J., Han, Y.S., Ko, S.K.: Decidability of involution hypercodes. Theoretical Computer Science **550**(0), 90 – 99 (2014)
4. Constantinescu, S., Ilie, L.: Fine and Wilf's theorem for Abelian periods. Bulletin of the EATCS **89**, 167–170 (2006)
5. Crochemore, M., Hancart, C., Lecroq, T.: Algorithms on Strings. Cambridge University Press (2007)
6. Crochemore, M., Rytter, W.: Jewels of Stringology. World Scientific (2002)
7. Cummings, L.J., Smyth, W.F.: Weak repetitions in strings. J. Combinatorial Mathematics and Combinatorial Computing **24**, 33–48 (1997)
8. Czeizler, E., Kari, L., Seki, S.: On a special class of primitive words. Theoretical Computer Science **411**, 617 – 630 (2010)
9. de Luca, A., De Luca, A.: Pseudopalindrome closure operators in free monoids. Theoretical Computer Science **362**(13), 282 – 300 (2006)
10. Gawrychowski, P., Manea, F., Mercaş, R., Nowotka, D., Tiseanu, C.: Finding pseudo-repetitions. Leibniz International Proceedings in Informatics **20**, 257–268 (2013)
11. Gawrychowski, P., Manea, F., Nowotka, D.: Discovering hidden repetitions in words. In: P. Bonizzoni, V. Brattka, B. Löwe (eds.) The Nature of Computation. Logic, Algorithms, Applications, *Lecture Notes in Computer Science*, vol. 7921, pp. 210–219. Springer Berlin Heidelberg (2013)
12. Gawrychowski, P., Manea, F., Nowotka, D.: Testing generalised freeness of words. In: E.W. Mayr, N. Portier (eds.) 31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), *Leibniz International Proceedings in Informatics (LIPIcs)*, vol. 25, pp. 337–349. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2014)
13. Hopcroft, J.E., Ullman, J.D.: Formal Languages and Their Relation to Automata. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (1969)
14. Hsu, S., Ito, M., Shyr, H.: Some properties of overlapping order and related languages. Soochow Journal of Mathematics **15**(1), 29–45 (1989)
15. Huang, C., Hsiao, P.C., Liau, C.J.: A note of involutively bordered words. Journal of Information and Optimization Sciences **31**(2), 371–386 (2010)

16. Hussini, S., Kari, L., Konstantinidis, S.: Coding properties of DNA languages. In: N. Jonoska, N. Seeman (eds.) Proc. of DNA7, *Lecture Notes in Computer Science*, vol. 2340, pp. 57–69. Springer (2002)
17. Jonoska, N., Kephart, D., Mahalingam, K.: Generating DNA code words. In: GECCO Late Breaking Papers, pp. 240–246 (2002)
18. Kari, L., Kulkarni, M.S.: Pseudo-identities and bordered words. In: G. Paun, G. Rozenberg, A. Salomaa (eds.) Discrete Mathematics and Computer Science, pp. 207–222. Editura Academiei Române (2014)
19. Kari, L., Mahalingam, K.: Involutively bordered words. International Journal of Foundations of Computer Science **18**(05), 1089–1106 (2007)
20. Kari, L., Mahalingam, K.: Watson-Crick conjugate and commutative words. In: M. Garzon, H. Yan (eds.) DNA Computing, *Lecture Notes in Computer Science*, vol. 4848, pp. 273–283. Springer Berlin Heidelberg (2008)
21. Kari, L., Mahalingam, K.: Watson-Crick palindromes in DNA computing. Natural Computing **9**(2), 297–316 (2010)
22. Kari, L., Seki, S.: On pseudoknot-bordered words and their properties. Journal of Computer and System Sciences **75**, 113 – 121 (2009)
23. Kari, L., Seki, S.: An improved bound for an extension of Fine and Wilf's theorem and its optimality. Fundamenta Informaticae **101**, 215–236 (2010)
24. Kopecki, S.: Personal communication (2015)
25. Lothaire, M.: Combinatorics on Words. Cambridge University Press (1997)
26. Shyr, H., Thierrin, G.: Disjunctive languages and codes. In: M. Karpinski (ed.) Fundamentals of Computation Theory, *Lecture Notes in Computer Science*, vol. 56, pp. 171–176. Springer Berlin Heidelberg (1977)
27. Shyr, H.J.: Free Monoids and Languages. Department of Mathematics, Soochow University, Taipei, Taiwan (1979)
28. Trappe, W., Washington, L.C.: Introduction to Cryptography with Coding Theory. Pearson Education India (2006)
29. Watson, J.D., Crick, F.H.: Molecular structure of nucleic acids. Nature **171**(4356), 737–738 (1953)
30. Wong, F.R.: Algebraic Properties of d-Primitive Words. Master's thesis, Chuang-Yuan Christian University, Chuang Li, Taiwan (1994)
31. Yu, S.: d-minimal languages. Discrete Applied Mathematics **89**(13), 243–262 (1998)
32. Yu, S.S.: Languages and Codes. Tsang Hai Book Publishing Co (2005)
33. Ziv, J., Lempel, A.: A universal algorithm for sequential data compression. IEEE Transactions on Information Theory **23**(3), 337–343 (1977)