# On the regularity of iterated hairpin completion of a single word

Lila Kari[*], Steffen Kopecki[†], and Shinnosuke Seki[‡] [§]

June 12, 2011

**Abstract**

Hairpin completion is an abstract operation modeling a DNA bio-operation which receives as input a DNA strand $w = x\alpha y\overline{\alpha}$, and outputs $w' = x\alpha y\bar{\alpha}\overline{x}$, where $\overline{x}$ denotes the Watson-Crick complement of $x$. In this paper, we focus on the problem of finding conditions under which the iterated hairpin completion of a given word is regular. According to the numbers of words $\alpha$ and $\overline{\alpha}$ that initiate hairpin completion and how they are scattered, we classify the set of all words $w$. For some basic classes of words $w$ containing small numbers of occurrences of $\alpha$ and $\overline{\alpha}$, we prove that the iterated hairpin completion of $w$ is regular. For other classes with higher numbers of occurrences of $\alpha$ and $\overline{\alpha}$, we prove a necessary and sufficient condition for the iterated hairpin completion of a word in these classes to be regular.

## 1 Introduction

A DNA strand can be abstractly viewed as a word over the alphabet $\{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{T}\}$, where in A is Watson-Crick complementary to T and C to G, and two complementary DNA single strands of opposite orientation bind together to form a double DNA strand (intermolecular structure). Also, if subwords of a DNA strand are complementary, the strand may bind to itself forming intramolecular structures such as stem-loops, also known more commonly as *hairpins* (Figure 1 (2)). Hairpins can be a building block of a larger-scale structure of RNA strands, and play a role in determining various chemical and thermodynamical properties (stability, functions) of the structure, and make significant contributions to

[*]Department of Computer Science University of Western Ontario London, Ontario, N6A 5B7, Canada, lila@csd.uwo.ca

[†]Institute for Formal Methods in Computer Science (FMI) University of Stuttgart, Universitätsstraße, D-70569, Stuttgart, Germany, kopecki@fmi.uni-stuttgart.de

[‡]Department of Systems Biosciences for Drug Discovery, Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29, Yoshida-Shimo-Adachi-cho, Sakyo-ku, Kyoto, 606-8501, Japan, sseki@pharm.kyoto-u.ac.jp

[§]corresponding author

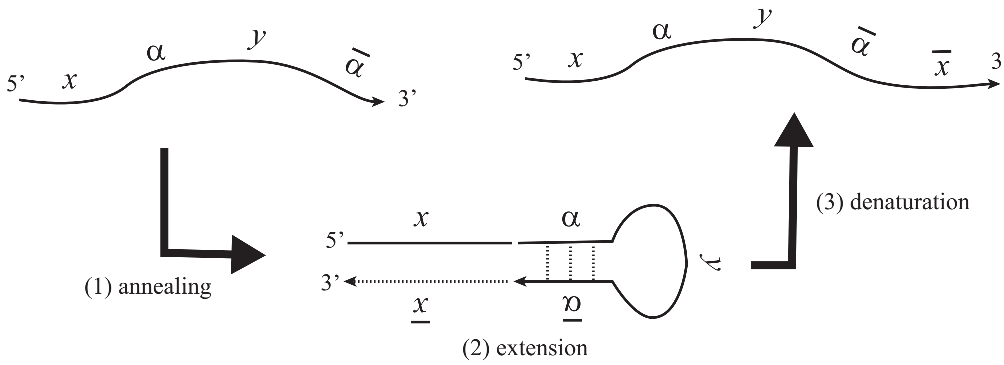


Figure 1: Hairpin completion by polymerase chain reaction. The operation input is $x\alpha y\overline{\alpha}$, the output is $x\alpha y\bar{\alpha}\overline{x}$, and the primer is $\alpha$.

the genetic information processing as illustrated in their function as a stopper for messenger RNA (mRNA) transcription. A `CG`-rich sequence of an mRNA folds into its Watson-Crick complement on the RNA and forms a stable hairpin. Transcription of the mRNA is terminated when RNA polymerase reaches the hairpin. At that time, *nusA* protein bound to the polymerase interacts with the hairpin and takes the polymerase off the mRNA. This hairpin-driven mechanism is called *intrinsic termination* [21]. As such, hairpins tend to interfere with reactions, and therefore were given the cold shoulder by DNA computing experimentalists. See [1, 2, 9, 10, 17] about this problem and about some of the "good" designs of DNA strands that are free of hairpins.

Hairpin is not a foe to all DNA computing experiments; many molecular computing machineries have been proposed which make good use of hairpins. Such hairpin-driven systems include DNA RAM [11, 19, 20] and Whiplash PCR [7, 18]. In particular, Whiplash PCR features a self-directed polymerase chain reaction (PCR) of DNA strand, which practically motivates the investigation of a formal language operation called *hairpin completion*. Hairpin completion proceeds as follows (Figure 1): Starting from a DNA strand $w = x\alpha y\overline{\alpha}$, a segment $\overline{\alpha}$ at the 3'-end of $w$ binds to its Watson-Crick complementary strand $\alpha$ on the strand (annealing). A polymerase chain reaction then extends $w$ at its 3'-end in the $5' \to 3'$ direction so as to generate the strand $x\alpha y\bar{\alpha}\overline{x}$ (let us call $\alpha$ and $\overline{\alpha}$ that bind with each other to initiate this PCR reaction *primers*). Despite the intrinsic $5' \to 3'$ polarity of polymerases, a mechanism exists to make polymerase reaction work in the $3' \to 5'$ direction (Okazaki fragment [16]).

As an abstract model of the above-mentioned self-directed PCR, Cheptea, Martín-Vide, and Mitrana proposed the hairpin completion in [3], and since then this abstract operation has been studied on its algorithmic and formal linguistic aspects [5, 13, 14, 15] together with its variant called *bounded hairpin* *completion* [8, 12], where the length of extension in one operation is bounded by a constant. Ito et al. [8] and Kopecki [12] proved that all classes in the Chomsky Hierarchy are closed under iterated bounded hairpin completion. In contrast, the class of regular languages was proved not to be closed under iterated (unbounded) hairpin completion [3], and a surprising fact is that iterated hairpin completion of a *single* word can be non-regular [12]. In this paper, we focus on a problem proposed by Kopecki in [12]; is it decidable whether the iterated hairpin completion of a given word is regular? The iterated hairpin completion of a singleton language (a word) is known to be in $\mathtt{NL}$ [3], but can be non-regular as shown in the following example.

*Example* 1. Let $\alpha = a^k$ and $w = \alpha b\alpha c\alpha\bar{\alpha}\bar{d}\bar{\alpha}$, where $a, \bar{a}, b, \bar{b}, c, \bar{c}, d, \bar{d}$ are all distinct letters. Then the intersection of the iterated hairpin completion of $w$ with $(\alpha b\alpha c(\alpha b)^+\alpha d)^2\alpha b\alpha c\alpha\bar{\alpha}\bar{d}\bar{\alpha}(\bar{b}\bar{\alpha})^+\bar{c}\bar{\alpha}\bar{b}\bar{\alpha}$ is $\{(\alpha b\alpha c(\alpha b)^i\alpha d)^2\alpha b\alpha c\alpha\bar{\alpha}\bar{d}\bar{\alpha}(\bar{b}\bar{\alpha})^i\bar{c}\bar{\alpha}\bar{b}\bar{\alpha} \mid i \geq 1\}$. This intersection is not context-free, and therefore, neither is the iterated hairpin completion.

In this paper, we give a partial answer to the regularity-test decidability problem. We focus our attention on the number of primers occurring on a given word as its factors and on how these primers are scattered over the given word. All the words are classified in accordance with these two criteria, and for some basic classes, we give a necessary and sufficient condition for the iterated hairpin completion of a word in the class to be regular.

# 2 Preliminaries

Let $\Sigma$ be an alphabet, $\Sigma^*$ be the set of all words over $\Sigma$, and for an integer $k \geq 0$, $\Sigma^k$ be the set of all words of length $k$ over $\Sigma$. The word of length 0 is called the empty word, denoted by $\lambda$, and let $\Sigma^+ = \Sigma^* \setminus \{\lambda\}$. A subset of $\Sigma^*$ is called a language over $\Sigma$. For a word $w \in \Sigma^*$, we employ the notation $w$ when we mean the word as well as the singleton language $\{w\}$ unless confusion arises. For a language $L \subseteq \Sigma^*$, we denote by $L^*$ the set $\{w_1 \cdots w_n \mid n \geq 0, w_1, \ldots, w_n \in L\}$.

We equip $\Sigma$ with a function $\bar{} : \Sigma \to \Sigma$ satisfying $\forall a \in \Sigma, \overline{\overline{a}} = a$; such a function is called an *involution*. This involution $\bar{}$ is naturally extended to words as: for $a_1, a_2, \ldots, a_n \in \Sigma$, $\overline{a_1 a_2 \cdots a_n} = \overline{a_n} \cdots \overline{a_2}\ \overline{a_1}$. For example, over the 4-letter alphabet $\Delta = \{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{T}\}$, if we define an involution $d : \Delta \to \Delta$ as $d(\mathtt{A}) = \mathtt{T}$, $d(\mathtt{T}) = \mathtt{A}$, $d(\mathtt{C}) = \mathtt{G}$, and $d(\mathtt{G}) = \mathtt{C}$, then $d$, being thus extended, maps the Watson strand of a complete DNA double strand into its Crick strand. For a word $w \in \Sigma^*$, we call $\overline{w}$ the *complement of* $w$, being inspired by this application. Another example is the *mirror image*, which is an involution $\mathrm{mi} : \Sigma^* \to \Sigma^*$ defined as $\mathrm{mi}(a) = a$ for all $a \in \Sigma$ and extended as such. For a language $L \subseteq \Sigma^*$, $\overline{L} = \{\overline{w} \mid w \in L\}$.

For words $u, w \in \Sigma^*$, if $w = xuy$ holds for some words $x, y \in \Sigma^*$, then $u$ is called a *factor* of $w$; a factor of $w$ that is strictly shorter than $w$ is said to be *proper*. If the equation holds with $x = \lambda$ ($y = \lambda$), then the factor $u$ is especially called a *prefix* (resp. a *suffix*) of $w$. The prefix relation can be regarded as a partial order $\leq_p$ over $\Sigma^*$; $u \leq_p w$ means that $u$ is a prefix of $w$. Analogously, by $w \geq_s v$ we mean that $v$ is a suffix of $w$. For a word $w \in \Sigma^*$ and a language $L \subseteq \Sigma^*$, a factor $v$ of $w$ is *minimal with respect to* $L$ if $v \in L$ and none of the proper factors of $v$ is in $L$.

A nonempty word $w \in \Sigma^+$ is *primitive* if $w = x^i$ implies $i = 1$ for any nonempty word $x \in \Sigma^+$. It is well-known that for any nonempty word $w$, there exists a unique primitive word $u$ with $w \in u^+$. Such $u$ is called the *primitive root of* $w$ and denoted by $\rho(w)$. Two words $x, y \in \Sigma^*$ *commute* if $xy = yx$, and this is known to be equivalent to $\rho(x) = \rho(y)$. See [4] for details of primitivity and commutativity of words and related results.

Now we introduce the operation investigated in this paper, that is, hairpin completion, and define it formally. Imagine that we have a DNA sequence $5' - \mathtt{CAATCGTATGAT} - 3'$. The suffix `GAT` can find its $d$-image as a factor `ATC` on this sequence. Hence, this DNA sequence may bend over into a hairpin form by `GAT` binding with `ATC`. This formation of hairpin structure leaves `CA` as a free sticky-end, and DNA polymerase converts it into the complete double strand by extending its 3'-end by $\mathtt{TG} = d(\mathtt{CA})$. This exemplifies the mechanism of hairpin completion. We call two words whose thus binding initiate hairpin completion *primers*. In the above example, `GAT` and `ATC` are primers.

An important fact is that primers must be so long that the formed partial hairpin is thermodynamically stable enough to serve as a scaffold for the extension phase. In fact, primers used in natural DNA synthesis or in PCR are usually of length 10-30 nucleotides. Throughout this paper, the primer length is fixed to be a constant $k$. Let $\alpha \in \Sigma^k$ be a primer. If a given word $w \in \Sigma^*$ has a factorization $u\alpha v\overline{\alpha}$ for some $u, v \in \Sigma^*$ and $\alpha \in \Sigma^k$, then its *right hairpin completion with respect to* $\alpha$ results in the word $u\alpha v\overline{\alpha}\overline{u}$. As long as $\alpha$ is clear from context, this operation is simply called *(single-primer) right hairpin completion*. By $w \to_{\mathcal{RH}_\alpha} w'$, or by $w \to_{\mathcal{RH}} w'$, we mean that $w'$ can be obtained from $w$ by right hairpin completion (with respect to $\alpha$). The *left hairpin completion* is defined analogously as an operation to derive $u'\alpha v'\overline{\alpha}\overline{u'}$ from $\alpha v'\overline{\alpha}\overline{u'}$, and the relation $\to_{\mathcal{LH}_\alpha}$ is naturally introduced. By $\to^*_{\mathcal{LH}}$ and $\to^*_{\mathcal{RH}}$, we denote the reflexive transitive closure of $\to_{\mathcal{LH}}$ and that of $\to_{\mathcal{RH}}$, respectively. The relation $\to_{\mathcal{H}}$ is defined as the union of $\to_{\mathcal{LH}}$ and $\to_{\mathcal{RH}}$.

For a given language $L \subseteq \Sigma^*$, we define the set of words obtained by left hairpin completion from $L$, and the set of words obtained by iterated left hairpin completion from $L$, respectively, as follows:

$$\mathcal{LH}_\alpha(L) = \{w' \mid \exists w \in L, w \to_{\mathcal{LH}_\alpha} w'\}, \qquad \mathcal{LH}^*_\alpha(L) = \{w' \mid \exists w \in L, w \to^*_{\mathcal{LH}_\alpha} w'\}.$$

Analogously, $\mathcal{RH}_\alpha(L)$ and $\mathcal{RH}^*_\alpha(L)$ are defined based on $\to_{\mathcal{RH}}$ and $\to^*_{\mathcal{RH}}$, and $\mathcal{H}_\alpha(L)$ and $\mathcal{H}^*_\alpha(L)$ are defined based on $\to_{\mathcal{H}}$ and $\to^*_{\mathcal{H}}$.

**Proposition 1.** *For a word* $w \in \Sigma^*$, $\mathcal{RH}^*_\alpha(w) = \overline{\mathcal{LH}^*_\alpha(\overline{w})}$.

Let us conclude the preliminaries by a brief investigation on the case when $\Sigma$ is unary. On such an alphabet $\Sigma = \{a\}$, there is only one involution definable, that is, $\bar{a} = a$ (identity). Hence, $\alpha = \overline{\alpha} = a^k$. By definition of hairpin completion, one can easily observe that for a word $w \in \Sigma^*$, if $|w| \leq 2k$, then $\mathcal{H}_\alpha^*(w) = \{w\}$. On the other hand, if $|w| > 2k$, then the prefix $\alpha = a^k$ (primer) of $w$ can bind with the second rightmost occurrence of $\overline{\alpha} = a^k$ on $w$, and hairpin completion extends $w$ to the left by $a$. This suggests that $\mathcal{H}_\alpha^*(w) = \{a^i \mid i \geq |w|\}$. In both cases, $\mathcal{H}_\alpha^*(w)$ is regular. Therefore, in the rest of this paper, we assume that $|\Sigma| \geq 2$.

# 3 Word structures relevant to the power of iterated hairpin completion

In this section, we describe several structural properties of a word $w$ that will be relevant for the characterization of its iterated hairpin completion $\mathcal{H}_\alpha^*(w)$, where $\alpha \in \Sigma^k$ is a fixed parameter.

## 3.1 Properties of $\alpha$-prefix, commutativity, and primitivity

A word $u \in \Sigma^*$ is called an *$\alpha$-prefix* of a word $w \in \Sigma^*$ if $u\alpha$ is a prefix of $w$. In a similar manner, a word $\overline{v} \in \Sigma^*$ is an *$\overline{\alpha}$-suffix* of $w$ if $\overline{\alpha}\,\overline{v}$ is a suffix of $w$. If $w = y\overline{\alpha}v$ begins with $\alpha$, then this prefix can bind with the occurrence of $\overline{\alpha}$ (unless they overlap with each other), and left hairpin completion results in $vw$. By $\mathrm{Pref}_\alpha(w)$ and $\mathrm{Suff}_{\overline{\alpha}}(w)$, we denote the set of all $\alpha$-prefixes and that of all $\overline{\alpha}$-suffixes of $w$, respectively. One can easily observe that $\mathrm{Suff}_{\overline{\alpha}}(w) = \overline{\mathrm{Pref}_\alpha(\overline{w})}$. Throughout this paper, we let $\mathrm{Pref}_\alpha(w) = \{u_0, \ldots, u_{m-1}\}$ and $\mathrm{Suff}_{\overline{\alpha}}(w) = \{\overline{v_0}, \ldots, \overline{v_{n-1}}\}$ for some $m, n \geq 0$. It will be convenient to assume that these $\alpha$-prefixes are sorted in the ascending order of their length. Likewise, we assume that $|\overline{v_0}| < |\overline{v_1}| < \cdots < |\overline{v_{n-1}}|$.

Our investigation on the properties of $\alpha$-prefix and $\overline{\alpha}$-suffix of a word begins with a basic observation.

**Proposition 2.** *For a word $w \in \alpha\Sigma^*$, the following statements hold: (1) for any $u \in \mathrm{Pref}_\alpha(w)$, $\alpha \leq_p u\alpha$; and (2) for any $x_1, \ldots, x_s \in \mathrm{Pref}_\alpha(w)$, $\alpha \leq_p x_1 \cdots x_s\alpha$;*

*Proof.* The first statement derives directly from the definition of $\alpha$-prefix and the assumption that $w$ begins with $\alpha$. For the second one, induction on $s$ works. Due to the first statement, $\alpha \leq_p x_s\alpha$ so that proving $\alpha \leq_p x_1 \cdots x_{s-1}x_s\alpha$ is reduced to proving $\alpha \leq_p x_1 \cdots x_{s-1}\alpha$. □

From this proposition, we can easily deduce that for a word $w \in \Sigma^*\overline{\alpha}$ and $\overline{y_1}, \ldots, \overline{y_t} \in \mathrm{Suff}_{\overline{\alpha}}(w)$, $\overline{\alpha}\,\overline{y_t} \cdots \overline{y_1} \geq_s \overline{\alpha}$, which means $\alpha \leq_p y_1 \cdots y_t\alpha$. This deepens the above observation further as follows.

**Corollary 3.** *For a word $w \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$, any word in $(\mathrm{Pref}_\alpha(w) \cup \overline{\mathrm{Suff}_{\overline{\alpha}}(w)})^*\alpha$ has $\alpha$ as its prefix.*

Due to the second statement of Proposition 2, $\alpha \leq_p x_1\alpha \leq_p x_1x_2\alpha \leq_p \cdots \leq_p x_1x_2 \cdots x_s\alpha$ holds for $\alpha$-prefixes $x_1, \ldots, x_s \in \mathrm{Pref}_\alpha(w)$. Hence, from a word $x_1x_2 \cdots x_s\alpha w'\overline{\alpha}$, one-step right hairpin completion can produce at least the words $x_1 x_2 \cdots x_s \alpha w' \overline{\alpha}\{\lambda, \overline{x_1}, \overline{x_1 x_2}, \ldots, \overline{x_1 x_2 \cdots x_s}\}$.[1] Now, if we know that one-step hairpin completion extends the word to the right by $\overline{u}$, what can we say about the word $u$? Firstly, as long as $|u| \leq |x_1 \cdots x_s|$, we can say that $u\alpha \leq_p x_1 \cdots x_s \alpha$ by definition of hairpin completion. Moreover, Corollary 3 enables us to find $0 \leq i < s$ such that $|x_1 \cdots x_i| < |u| \leq |x_1 \cdots x_{i+1}|$. Then, one can let $u = x_1 \cdots x_i z$ for some prefix $z$ of $x_{i+1}$. Since $z\alpha \leq_p x_{i+1}\alpha \leq_p w$, $z$ is an $\alpha$-prefix of $w$. By defining $\mathrm{ind}(x_{i+1})$ to be the index satisfying $u_{\mathrm{ind}(x_{i+1})} = x_{i+1}$, we have $z \in \{u_0, u_1, \ldots, u_{\mathrm{ind}(x_{i+1})}\}$ (recall that elements of $\mathrm{Pref}_\alpha(w)$ are sorted with respect to their length). Thus, the next lemma holds.

**Lemma 4.** *Let $x_1, \ldots, x_s \in \mathrm{Pref}_\alpha(w)$. If a word $u$ satisfies $u\alpha \leq_p x_1 \cdots x_s \alpha$, then there exists an integer $0 \leq i < s$ such that $u = x_1 \cdots x_i z$ for some $z \in \{u_0, u_1, \ldots, u_{\mathrm{ind}(x_{i+1})}\}$.*

A more natural setting is that each of $x_1, \ldots, x_s$ be an element of either $\mathrm{Pref}_\alpha(w)$ *or* $\overline{\mathrm{Suff}_{\overline{\alpha}}(w)}$ for left hairpin completion can produce the complement of a $\overline{\alpha}$-suffix of $w$ to the left of $w$. We need to generalize the function ind by extending its domain as: if $x_i = v_j \in \overline{\mathrm{Suff}_{\overline{\alpha}}(w)}$, then $\mathrm{ind}(x_i) = j$.

**Lemma 5.** *Let $x_1, \ldots, x_s \in \mathrm{Pref}_\alpha(w) \cup \overline{\mathrm{Suff}_{\overline{\alpha}}(w)}$. If a word $u$ satisfies $u\alpha \leq_p x_1 \cdots x_s \alpha$, then there exists an integer $0 \leq i < s$ such that $u = x_1 \cdots x_i z$, where*

$$\begin{cases} z \in \{u_0, u_1, \ldots, u_{\mathrm{ind}(x_{i+1})}\} & \text{if } x_{i+1} \in \mathrm{Pref}_\alpha(w); \\ z \in \{v_0, v_1, \ldots, v_{\mathrm{ind}(x_{i+1})}\} & \text{if } x_{i+1} \in \overline{\mathrm{Suff}_{\overline{\alpha}}(w)}. \end{cases}$$

*Proof.* As done previously, we can find $0 \leq i < s$ and a nonempty word $z \in \Sigma^+$ satisfying $u = x_1 \cdots x_i z$ and $z\alpha \leq_p x_{i+1}\alpha$. Since this prefix relation can be rewritten as $\bar{\alpha}\overline{x_{i+1}} \geq_s \bar{\alpha}\overline{z}$, if $\overline{x_{i+1}}$ is an $\overline{\alpha}$-suffix of $w$, so is $\overline{z}$. The case when $x_{i+1} \in \mathrm{Pref}_\alpha(w)$ is clear from the previous argument. □

Having considered prefix relations among $\alpha$-prefixes and $\overline{\alpha}$-suffixes of a word, now we proceed our study to more general factor relationships among them.

**Lemma 6.** *Let $w \in \alpha\Sigma^*$. If $u_j\alpha \geq_s u_i\alpha$ for some $1 \leq i < j < m$, then $u_j \in \{u_0, u_1, \ldots, u_{j-1}\}u_i$.*

*Proof.* We can let $xu_i\alpha = u_j\alpha$ for some $x \in \Sigma^*$. Combining this with Proposition 2, we have $x\alpha \leq_p u_j\alpha$ so that $x \in \mathrm{Pref}_\alpha(w)$. Since $|x| < |u_j|$, $x$ is one of $u_0, u_1, \ldots, u_{j-1}$. □

**Lemma 7.** *Let $w \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$. If $v_1\alpha$ is a factor of $u_1\alpha$, then $u_1 = v_1$.*

*Proof.* Due to the assumption on $w$, $u_0 = \overline{v_0} = \lambda$, and hence, $u_1$ and $\overline{v_1}$ are the shortest nonempty $\alpha$-prefix and $\overline{\alpha}$-suffix of $w$, respectively. Let $u_1\alpha = xv_1\alpha y$ for some $x, y \in \Sigma^*$. Unless $y = \lambda$, from $xv_1\alpha \leq_p u_1\alpha$, the word $xv_1$ would be a nonempty $\alpha$-prefix of $w$ that is strictly shorter than $u_1$, a contradiction. Thus, $y$ must be empty so that $u_1\alpha = xv_1\alpha$. Now, Lemma 6 leads us to $x = u_0 = \lambda$. □

[1] $\overline{x_1 x_2 \cdots x_s} = \overline{x_s} \cdots \overline{x_2}\ \overline{x_1}$.

Finally, let us introduce interesting results that illustrate the close relationship between $\alpha$-prefixes, commutativity, and primitivity, essential notions in combinatorics on words.

**Lemma 8.** *Let $w \in \alpha\Sigma^*$ and $u \in \mathrm{Pref}_\alpha(w)$. Then $\rho(u), \rho(u)^2, \ldots, \rho(u)^{|u|/|\rho(u)|} \in \mathrm{Pref}_\alpha(w)$.*

*Proof.* Due to the first statement of Proposition 2, $u \in \mathrm{Pref}_\alpha(w)$ enables us to let $\alpha y = u\alpha$ for some $y \in \Sigma^+$. Its solution is well-known to be $u = (st)^n$ and $\alpha = (st)^i s$ for some $n, i \geq 0$ and $s, t \in \Sigma^*$ such that $\rho(u) = st$. Hence, $u\alpha = (st)^{i+n}s = \rho(u)\alpha(ts)^{n-1} = \rho(u)^2\alpha(ts)^{n-2} = \cdots = \rho(u)^n\alpha$. $\square$

An immediate implication of this lemma is that the shortest nonempty $\alpha$-prefix of a word that begins with $\alpha$ must be primitive. We should make one more step forward. Imagine that a word $w$ has an $\alpha$-prefix $u$. If $w \to_{\mathcal{RH}} w\overline{u}$ is possible, then $w \to_{\mathcal{RH}} w\overline{\rho(u)}$ is also possible. Thus, repeating the extension of $w$ by $\overline{\rho(u)}$ to the right $|u|/|\rho(u)|$ times amounts to extending $w$ by $\overline{u}$ once. In other words, the process to extend a word by $\overline{u}$ is not essential unless $u$ is primitive because it can be always simulated by multiple processes to extend a word by $\overline{\rho(u)}$.

The next lemma proves that all nonempty $\alpha$-prefixes of length at most $|\alpha|$ commute with each other, and hence, only the shortest one is essential in the above sense.

**Lemma 9.** *For nonempty words $x_1, x_2 \in \Sigma^+$, if $\alpha \leq_p x_1\alpha \leq_p x_2\alpha$ and $|x_2| \leq |\alpha|$ hold, then $\rho(x_1) = \rho(x_2)$.*

*Proof.* It suffices to consider the case when $|x_1| < |x_2|$. Combining $|x_1| \leq |\alpha|$ with $\alpha \leq_p x_1\alpha$, we can deduce that the word $x_1\alpha$ has a period $|x_1|$. Likewise, $x_2\alpha$ has a period $|x_2|$, and hence, $x_1\alpha$ also has this period. As a result, $x_1\alpha$ has two periods $|x_1|, |x_2|$, and moreover it is of length at least the sum of these periods. Thus, Fine and Wilf's theorem [4, 6] leads us to the conclusion of this lemma. $\square$

## 3.2 Non-crossing words and their properties

A word $w_0 \in \Sigma^*$ is an $(m,n)$-*$\alpha$-word*, or simply an $(m,n)$-*word* when $\alpha$ is clear from the context, if $|\mathrm{Pref}_\alpha(w_0)| = m$ and $|\mathrm{Suff}_{\overline{\alpha}}(w_0)| = n$. Informally speaking, an $(m,n)$-word is a word on which $\alpha$ occurs $m$ times and $\overline{\alpha}$ does $n$ times. If $\alpha = \overline{\alpha}$, then we regard an occurrence of $\alpha$ also as that of $\overline{\alpha}$, and as such, any word is an $(m,m)$-word for some $m \geq 0$.

We say that $w_0$ is *non-$\alpha$-crossing* if the rightmost occurrence of $\alpha$ precedes the leftmost one of $\overline{\alpha}$ on $w_0$. When $\alpha$ is understood from the context, we simply say that $w_0$ is non-crossing. Otherwise, the word is *$\alpha$-crossing* or *crossing*. Note that if $\alpha = \overline{\alpha}$, then for a word $w$ which is either a $(0,0)$-word or $(1,1)$-word, $\mathcal{H}^*_\alpha(w) = \{w\}$, and otherwise ($w$ is an $(m,m)$-word for some $m \geq 2$), $w$ can be considered crossing. Thus, whenever the non-$\alpha$-crossing word is concerned, we assume that $\alpha \neq \overline{\alpha}$. The definition of a word being non-$\alpha$-crossing does not force the word to begin with $\alpha$ or end with $\overline{\alpha}$. However, it is not until $\alpha$ is a primer that this notion becomes useful in our work. Thus, the word should be in either $\alpha\Sigma^*$ or $\Sigma^*\overline{\alpha}$. Actually, in the rest of this paper, we assume both of these conditions and consider only *single-primer iterated hairpin completion*; thus, we can assume that $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$. As let previously, elements of $\mathrm{Pref}_\alpha(w_0)$ are denoted by $u_0, u_1, \ldots, u_{m-1}$, those of $\mathrm{Suff}_{\overline{\alpha}}(w_0)$ by $\overline{v_0}, \overline{v_1}, \ldots, \overline{v_{n-1}}$, and they are sorted so that this assumption imposes $u_0 = v_0 = \lambda$.

Our main focus lies on the characterization of non-crossing words whose iterated hairpin completion is regular in terms of combinatorics on words. Thus, in this subsection, we prove some combinatorial properties of non-crossing words.

Let us begin with two basic observations: the first one is about the longest $\alpha$-prefix and $\overline{\alpha}$-suffix of $w_0$ and the second one is about the closure property of a word being non-crossing under hairpin completion. Especially, this closure property forms the foundation of our discussions in the rest of this paper.

**Proposition 10.** *$u_{m-1} = v_{n-1}$ if and only if $m = n$ and for all $0 \leq i < m$, $u_i = v_i$.*

**Proposition 11.** *Let $\alpha \in \Sigma^k$ with $\alpha \neq \overline{\alpha}$, and $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ be a non-crossing word. Then any word in $\mathcal{H}_\alpha^*(w_0)$ is non-crossing.*

By definition, hairpin completion can extend $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ to the right by $\overline{u_i}$ for some $0 \leq i < m$ if and only if $|u_i| + 2|\alpha| \leq |w_0|$, i.e., the occurrence of $\alpha$ just to the right of $u_i$ on $w_0$ does not overlap with the suffix $\overline{\alpha}$ of $w_0$. Thus, if $w_0 \to_{\mathcal{RH}} w_0\overline{u_i}$ is valid, then $w_0 \to_{\mathcal{RH}} w_0\overline{u_j}$ is also valid for any $0 \leq j \leq i$. These observations motivate us to ask the question of whether $w_0 \to_{\mathcal{RH}} w_0\overline{u_{m-1}}$ or $w_0 \to_{\mathcal{LH}} v_{n-1}w_0$ is always valid.

**Lemma 12.** *Let $w_0$ be a non-crossing $(m,n)$-word with $\mathrm{Pref}_\alpha(w_0) = \{u_0, \ldots, u_{m-1}\}$ and $\mathrm{Suff}_{\overline{\alpha}}(w_0) = \{\overline{v_0}, \ldots, \overline{v_{n-1}}\}$. Then $|u_{m-2}| + |v_{n-1}| + 2|\alpha| < |w_0|$.*

*Proof.* Suppose that this inequality did not hold. Being non-crossing, $w_0$ can be written as $w_0 = u_{m-2}w\overline{v_{n-1}}$ for some $w \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ with $|w| \leq 2|\alpha|$; this length condition imposes $w = \overline{w}$. Let $x$ be a nonempty word satisfying $u_{m-1} = u_{m-2}x$. Since $w_0$ is non-crossing, $u_{m-1}\alpha \leq_p u_{m-2}w$ must hold, from which we have $x\alpha \leq_p w$. Combining this with $w = \overline{w}$ enables us to find an $\overline{\alpha}$-suffix $\overline{x}\,\overline{v_{n-1}}$ of $w_0$, but this would be strictly longer than the longest $\overline{\alpha}$-suffix of $w_0$, a contradiction. □

Thus, $w_0 \to_{\mathcal{RH}} w_0\overline{u_{m-2}}$ and $w_0 \to_{\mathcal{LH}} v_{n-2}w_0$ are valid, and so are $w_0 \to_{\mathcal{RH}} w_0\overline{u_i}$ and $w_0 \to_{\mathcal{LH}} v_jw_0$ for $0 \leq i \leq m-2$ and $0 \leq j \leq n-2$. This lemma does not rule out the possibility that $w_0$ cannot be extended to the right by $\overline{u_{m-1}}$ by hairpin completion. This is because the occurrence of $\alpha$ to the right of $u_{m-1}$ might overlap with the suffix $\overline{\alpha}$. The analogous argument is valid for $v_{n-1}$ and left hairpin completion. However, Lemma 12 leads us to one important corollary on non-crossing $(m,n)$-words for $m, n \geq 2$ that hairpin completion can extend $w_0$ to the right by the complement of any of its $\alpha$-prefix and to the left by the complement of any of its $\overline{\alpha}$-suffix.

**Corollary 13.** *Let $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ be a non-crossing $(m,n)$-word with $m,n \geq 2$. Then $\mathcal{H}_\alpha(w_0) = \{w_0\} \cup \{v_0, v_1, \ldots, v_{m-1}\}w_0 \cup w_0\{\overline{u_0}, \overline{u_1}, \ldots, \overline{u_{n-1}}\}$.*

We conclude this section with a characterization of a non-$\alpha$-crossing word in terms of minimal factors with respect to the language $\alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$. With Proposition 11, this characterization will bring a unique factorization theorem (Theorem 15) of any word $w$ in $\mathcal{H}^*_\alpha(w_0)$ as $w = xw_0y$ for some words $x, y$, and furthermore, we can easily observe that $x$ is generated by left hairpin completions and $y$ is by right hairpin completions.

**Lemma 14.** *Let $\alpha \in \Sigma^k$ with $\alpha \neq \overline{\alpha}$. A word $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ is non-crossing if and only if it contains exactly one minimal factor $v$ from $\alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$.*

*Proof.* Let us consider the contrapositive of the converse implication. So, if $w_0$ is crossing, then we can find an occurrence of $\overline{\alpha}$ (let us denote it by $\overline{\alpha}_0$) which precedes an occurrence of $\alpha$ ($\alpha_1$). $\overline{\alpha}_0$ is guaranteed to be preceded by another occurrence of $\alpha$ ($\alpha_2$) because $w_0$ begins with $\alpha$. Thus, the factor of $w_0$ that spans from $\alpha_2$ to $\overline{\alpha}_0$ is a minimal factor from $\alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$. By the same token, the factor of $w_0$ that spans from $\alpha_1$ to its right adjacent occurrence of $\overline{\alpha}$ becomes another minimal factor.

In order to prove the direct implication, suppose that $w_0$ contains two minimal factors from $\alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$. These two factors must overlap with each other because otherwise the suffix $\overline{\alpha}$ of the first factor precedes the prefix $\alpha$ of the second one and $w$ would be crossing. However, if they overlap, then the overlapped part would be in $\alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$, and this contradicts the minimality of the two factors. ☐

**Theorem 15.** *Let $\alpha \in \Sigma^k$ with $\alpha \neq \overline{\alpha}$, and $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ be a non-crossing word. On any word in $\mathcal{H}^*_\alpha(w_0)$, $w_0$ occurs exactly once as a factor.*

*Proof.* This is from the two facts that any word in $\mathcal{H}^*_\alpha(w_0)$ is non-crossing (Proposition 11) and that these words contain at least one occurrence of $w_0$ as a factor by definition of hairpin completion. ☐

# 4 Iterated hairpin completion of non-crossing words

This section contains the main contribution of this paper: characterizations of the regularity of iterated hairpin completion of a non-crossing $(m,n)$-word $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ (recall that $\alpha \neq \overline{\alpha}$ is assumed). Throughout this section, $w_0$ is thus assumed with $\mathrm{Pref}_\alpha(w_0) = \{u_0, u_1, \ldots, u_{m-1}\}$ and $\mathrm{Suff}_{\overline{\alpha}}(w_0) = \{\overline{v_0}, \overline{v_1}, \ldots, \overline{v_{n-1}}\}$, where elements of each set is sorted in the order of their lengths; so $u_0 = v_0 = \lambda$.

Let us begin with a proof that one-sided hairpin completion of a non-crossing word is regular (Theorem 16). Then we will show that the iterated hairpin completion of a non-crossing $(m,1)$-word for any $m \geq 1$ or $(2,2)$-word is always regular (Theorems 17 and 19). Using these results and combinatorial results shown in Section 3, we characterize the set of all non-crossing $(3,2)$-words whose iterated hairpin completion is regular, in terms of commutativity (Theorem 24).

**Theorem 16.** *For a non-crossing word $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$, both $\mathcal{LH}^*_\alpha(w_0)$ and $\mathcal{RH}^*_\alpha(w_0)$ are regular.*

*Proof.* First, we prove the regularity of $\mathcal{RH}^*_\alpha(w_0)$. Let $w$ be an $\alpha$-prefix of $w_0$. A right hairpin completion of $w_0$ can produce $w_0\overline{w}$. Note that the suffix $\bar{\alpha}\overline{w}$ of this resulting word does not contain $\alpha$ due to the non-crossing assumption on $w_0$, and this means that the longest $\alpha$-prefix of $w_0\overline{w}$ is the same as that of $w_0$. Thus, the language $\mathcal{RH}^*_\alpha(w_0)$ can be obtained by iterated bounded hairpin completion from $w_0$, and hence, is regular [12].

For the regularity of $\mathcal{LH}^*_\alpha(w_0)$, it suffices to observe that $\overline{w_0}$ is also non-crossing. Using the result just proved, $\mathcal{RH}^*_\alpha(\overline{w_0})$ is regular, and according to Proposition 1, $\mathcal{LH}^*_\alpha(w_0) = \overline{\mathcal{RH}^*_\alpha(\overline{w_0})}$. Note that the class of regular languages is closed under $^-$. □

## 4.1 Iterated hairpin completion of non-crossing $(m,1)$-words

In this subsection, we consider the case $n = 1$ ($w_0$ is an $(m,1)$-word), and prove that $\mathcal{H}^*_\alpha(w_0)$ is regular. For $m = 1$, it is easy to see that hairpin completion cannot generate any word but $w_0$, that is, $\mathcal{H}^*_\alpha(w_0) = \{w_0\}$. Hence, we assume $m \geq 2$.

Lemma 12 means that hairpin completion can extend $w_0$ to the right by any of $\overline{u_0}, \overline{u_1}, \ldots, \overline{u_{m-2}}$, As mentioned previously, in contrast, $w_0$ is thus extendable by $\overline{u_{m-1}}$ only if $|u_{m-1}| + 2|\alpha| \leq |w_0|$. As a result, if $m = 2$ but this inequality does not hold, then $\mathcal{H}^*_\alpha(w_0) = \{w_0\}$. Hence, we can advance our discussion on the assumption that $w_0 \to_{\mathcal{RH}} w_0\overline{u_1}$ is valid.

Note that $w_0\overline{u_1}$ is a non-crossing $(m,2)$-word. Applying Lemma 12 to this word, we can see that $|u_{m-1}| + 2|\alpha| < |w_0\overline{u_1}|$. Hence, hairpin completion can extend $w_0\overline{u_1}$ further to the right not only by any of $\overline{u_0}, \overline{u_1}, \ldots, \overline{u_{m-2}}$ but also by $\overline{u_{m-1}}$. We can strengthen this observation that any of $w_0\overline{u_1}, \ldots, w_0\overline{u_{m-1}}$ is thus extendable.

Let us define the following regular language:

$$R_{m1}(w_0) = \{w_0\} \cup \Big\{x_s \cdots x_1 w_0 \overline{y_1}\,\overline{y_2} \cdots \overline{y_t} \;\Big|\; \begin{array}{l} y_1 \in \begin{cases} \{u_0, u_1, \ldots, u_{m-2}, u_{m-1}\} & \text{if } |u_{m-1}| + 2|\alpha| \leq |w_0| \\ \{u_0, u_1, \ldots, u_{m-2}\} & \text{otherwise} \end{cases} \\ s \geq 0, t \geq 1, x_s, \ldots, x_1, y_2, \ldots, y_t \in \{u_0, u_1, \ldots, u_{m-1}\}, \\ \text{and } \max_{1 \leq i \leq s}\{\mathrm{ind}(x_i)\} \leq \max_{1 \leq j \leq t}\{\mathrm{ind}(y_j)\} \end{array} \Big\}.$$

We claim that this language is the language obtained from $w_0$ by iterated hairpin completion.

First, we prove that $\mathcal{H}^*_\alpha(w_0) \supseteq R_{m1}(w_0)$. Let $w \in R_{m1}(w_0)$. By definition, any word in $R_{m1}(w_0)$ can be factorized as $w = x_s \cdots x_1 w_0 \overline{y_1}\,\overline{y_2} \cdots \overline{y_t}$. Compare the leftmost factor $x_s$ to the complement of the rightmost factor $\overline{y_t}$ with respect to their index. Consider the case when $\mathrm{ind}(x_s) \leq \mathrm{ind}(y_t)$. Recall that this means $x_s = u_{\mathrm{ind}(x_s)}$ and $y_t = u_{\mathrm{ind}(y_t)}$, and that $u_{\mathrm{ind}(x_s)}\alpha \leq_p u_{\mathrm{ind}(v_t)}\alpha$. Thus, $w_0\overline{y_1}\cdots\overline{y_t} \geq_s \overline{\alpha}\,\overline{y_t} \geq_s \overline{\alpha}\,\overline{x_s}$. This means that $x_{s-1}\cdots x_1 w_0\overline{y_1}\cdots\overline{y_t} \to_{\mathcal{LH}} w$ is valid. By the same token, $x_s\cdots x_1 w_0\overline{y_1}\cdots\overline{y_{t-1}} \to_{\mathcal{RH}} w$ is valid when $\mathrm{ind}(x_s) > \mathrm{ind}(y_t)$. Due to $\max_{1\leq i\leq s}\{\mathrm{ind}(x_i)\} \leq \max_{1\leq j\leq t}\{\mathrm{ind}(y_j)\}$, the repetition of this process eventually leads us to a word $w_0\overline{y_1}\cdots\overline{y_j}$ for some $1 \leq j \leq t$ such that $w_0\overline{y_1}\cdots\overline{y_j} \to^*_{\mathcal{H}} w$. Because of the condition on $y_1$ and our discussion above, $w_0 \to_{\mathcal{RH}} w_0\overline{y_1} \to_{\mathcal{RH}} \cdots \to_{\mathcal{RH}} w_0\overline{y_1}\cdots\overline{y_j}$ is valid so that $w \in \mathcal{H}^*_\alpha(w_0)$. Hence, $\mathcal{H}^*_\alpha(w_0) \supseteq R_{m1}(w_0)$.

Secondly, we prove the opposite inclusion by induction on the length of derivation by hairpin completion. Clearly $w_0 \in R_{m1}(w_0)$. Let us assume that a word in $\mathcal{H}^*_\alpha(w_0)$ can be written as $x_s\cdots x_1 w_0\overline{y_1}\cdots\overline{y_t}$ with $\max_{1\leq i\leq s}\{\mathrm{ind}(x_i)\} \leq \max_{1\leq j\leq t}\{\mathrm{ind}(y_j)\}$. Let $\ell = \max_{1\leq j\leq t}\{\mathrm{ind}(y_j)\}$. If hairpin completion extends this word to the left by $x$, then $\overline{\alpha}\,\overline{y_1}\cdots\overline{y_t} \geq_s \overline{\alpha}\,\overline{x}$ and this means $x \in \{u_0, u_1, \ldots, u_\ell\}^+$ (see Lemma 4). Thus, there exist some $s' > s$ and $x_{s'}, \ldots, x_{s+1} \in \{u_0, u_1, \ldots, u_\ell\}$ such that $x = x_{s'}\cdots x_{s+1}$ and $\max\{\mathrm{ind}(x_{s'}), \ldots, \mathrm{ind}(x_{s+1}), \mathrm{ind}(x_s), \ldots, \mathrm{ind}(x_1)\} \leq \ell$. It it trivial that this inequality remains valid in the right hairpin completion.

**Theorem 17.** *Let $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ be a non-crossing $(m,1)$-word for some $m \geq 1$. Then $\mathcal{H}^*_\alpha(w_0)$ is regular.*

The key idea in the above discussion is that if a word in $\mathcal{H}^*_\alpha(w_0)$ begins with the longest $\alpha$-prefix $u_{m-1}$ of $w_0$ and the word is of length at least $|u_{m-1}| + 2|\alpha|$, then hairpin completion can extend it to the right by the complement of any $\alpha$-prefix of $w_0$. This idea has a broader range of applications. Let $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ be a non-crossing $(m,n)$-word for some $m, n \geq 1$ with $\mathrm{Pref}_\alpha(w_0) = \{u_0, u_1, \ldots, u_{m-1}\}$ and $\mathrm{Suff}_{\overline{\alpha}}(w_0) = \{\overline{v_0}, \overline{v_1}, \ldots, \overline{v_{n-1}}\}$. Proposition 10 says that if $u_{m-1} = v_{n-1}$, then $\mathrm{Suff}_{\overline{\alpha}}(w_0) = \overline{\mathrm{Pref}_\alpha(w_0)}$. For $m \geq 2$, the rightmost occurrence of $\alpha$ on $w_0$ does not overlap with the suffix $\overline{\alpha}$ of $w_0$ (Lemma 12). Thus, $\mathcal{H}^*_\alpha(w_0) = \{u_0, \ldots, u_{m-1}\}^* w_0 \{\overline{u_0}, \ldots, \overline{u_{m-1}}\}^*$.

**Corollary 18.** *Let $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ be a non-crossing $(m,n)$-word. If $u_{m-1} = v_{n-1}$, then $\mathcal{H}^*_\alpha(w_0)$ is regular.*

### 4.2 Iterated hairpin completion of non-crossing $(2,2)$-words

In contrast to the result obtained in the previous subsection, Example 1 shows that there exists a non-crossing $(m,2)$-word whose iterated hairpin completion is non-regular with $m = 3$. This result motivates the study of non-crossing $(2,2)$-words reported here. Let $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ be a non-crossing $(2,2)$-word. We can employ Corollary 13 to see that $\mathcal{H}_\alpha(w_0) = \{w_0, v_1 w_0, w_0\overline{u_1}\}$, which means

$$\mathcal{H}^*_\alpha(w_0) = \{w_0\} \cup \mathcal{H}^*_\alpha(v_1 w_0) \cup \mathcal{H}^*_\alpha(w_0\overline{u_1}).$$

We will prove the regularity of the second and third terms of the above equation in order to prove that of $\mathcal{H}^*_\alpha(w_0)$. For this goal, it is useful to observe here that a word $xw_0\overline{y}$ in $\mathcal{H}^*_\alpha(w_0)$ can be extended to the right by the complement of any of the $\alpha$-prefixes of $x$ as well as $xu_0$ and $xu_1$, and to the left by the complement of any of the $\overline{\alpha}$-suffixes of $\overline{y}$ as well as $\overline{v_0}\,\overline{y}$ and $\overline{v_1}\,\overline{y}$.

Let $R_{22L} = v_1^*(v_1w_0)\overline{v_1}^* \cup (v_1^+u_1)^*v_1^*(v_1w_0)\overline{v_1}^*(\overline{u_1}\,\overline{v_1}^+)^+$. We will show that this language is exactly the set of words obtained by iterated hairpin completion from $v_1w_0$.

In order to prove that $\mathcal{H}_\alpha^*(v_1w_0) \supseteq R_{22L}$, it suffices to present the following process:

$$\begin{aligned}
v_1w_0 &\to_{\mathcal{RH}}^* v_1w_0\overline{v_1}^{j_0} \to_{\mathcal{RH}} v_1w_0\overline{v_1}^{j_0}\overline{u_1}\,\overline{v_1}\\
&\to_{\mathcal{RH}}^* v_1w_0\overline{v_1}^{j_0}\overline{u_1}\,\overline{v_1}^{j_1} \to_{\mathcal{RH}} v_1w_0\overline{v_1}^{j_0}\overline{u_1}\,\overline{v_1}^{j_1}\overline{u_1}\,\overline{v_1}\\
&\to_{\mathcal{RH}}^* v_1w_0\overline{v_1}^{j_0}\overline{u_1}\,\overline{v_1}^{j_1}\cdots\overline{u_1}\,\overline{v_1}^{j_{t-1}}\overline{u_1}\,\overline{v_1}\\
&\to_{\mathcal{LH}}^* v_1^{i_0}v_1w_0\overline{v_1}^{j_0}\overline{u_1}\,\overline{v_1}^{j_1}\cdots\overline{u_1}\,\overline{v_1}^{j_{t-1}}\overline{u_1}\,\overline{v_1} \to_{\mathcal{LH}} v_1u_1v_1^{i_0}v_1w_0\overline{v_1}^{j_0}\overline{u_1}\,\overline{v_1}^{j_1}\cdots\overline{u_1}\,\overline{v_1}^{j_{t-1}}\overline{u_1}\,\overline{v_1}\\
&\to_{\mathcal{LH}}^* v_1^{i_1}u_1v_1^{i_0}v_1w_0\overline{v_1}^{j_0}\overline{u_1}\,\overline{v_1}^{j_1}\cdots\overline{u_1}\,\overline{v_1}^{j_{t-1}}\overline{u_1}\,\overline{v_1}\\
&\to_{\mathcal{LH}}^* v_1^{i_s}u_1\cdots v_1^{i_1}u_1v_1^{i_0}v_1w_0\overline{v_1}^{j_0}\overline{u_1}\,\overline{v_1}^{j_1}\cdots\overline{u_1}\,\overline{v_1}^{j_{t-1}}\overline{u_1}\,\overline{v_1}\\
&\to_{\mathcal{RH}}^* v_1^{i_s}u_1\cdots v_1^{i_1}u_1v_1^{i_0}v_1w_0\overline{v_1}^{j_0}\overline{u_1}\,\overline{v_1}^{j_1}\cdots\overline{u_1}\,\overline{v_1}^{j_{t-1}}\overline{u_1}\,\overline{v_1}^{j_t}.
\end{aligned}$$

Next, we prove the opposite inclusion by induction on the length of derivation by hairpin completion from $v_1w_0$. Obviously, $v_1w_0 \subseteq R_{22L}$. Assume that all words obtained from $v_1w_0$ by at most $i$ hairpin completions are in $R_{22L}$. Let $w_i$ be such a word and consider a word $w_{i+1}$ such that $w_i \to_{\mathcal{H}} w_{i+1}$. Consider the case when this is a right hairpin completion. The rightmost occurrence of $\alpha$ on $w_i$ is the second $\alpha$ on its (unique) factor $w_0$. Therefore, if we let $w_{i+1} = w_i\overline{x}$, then $x\alpha \leq_p (v_1^+u_1)^*v_1^*v_1u_1\alpha$. Since $u_1$ and $\overline{v_1}$ are the respective shortest nonempty $\alpha$-prefix and $\overline{\alpha}$-suffix of $w_0$, Lemma 5 implies that $x \in \overline{(v_1^+u_1)^*v_1^*}$. Note that $R_{22L}$ is closed under catenating a word in $\overline{(v_1^+u_1)^*v_1^*}$ to the right. Thus, $w_{i+1} \in R_{22L}$. The case when $w_i \to_{\mathcal{LH}} w_{i+1}$ can be proved in a symmetric manner.

Due to the symmetry of $u_1$ and $\overline{v_1}$, we can easily construct a regular language $R_{22R}$ which is equivalent to $H_\alpha^*(w_0\overline{u_1})$. Now the regularity of $\mathcal{H}_\alpha^*(w_0)$ has been proved.

**Theorem 19.** *If $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ is a non-crossing $(2,2)$-word, then $\mathcal{H}_\alpha^*(w_0)$ is regular.*

## 4.3 Iterated hairpin completion of non-crossing $(3,2)$-words

Theorem 19 and Example 1 motivate our investigation of non-crossing $(3,2)$-words. Actually, Theorem 24, our main contribution, provides a characterization of the regularity of iterated hairpin completion of a non-crossing $(3,2)$-word in terms of the commutativity of the $\alpha$-prefixes and $\overline{\alpha}$-suffixes of the word.

Let $w_0 \in \alpha\Sigma^*\cap\Sigma^*\overline{\alpha}$ be a non-crossing $(3,2)$-word (so $\alpha \neq \overline{\alpha}$) with $\mathrm{Pref}_\alpha(w_0) = \{u_0, u_1, u_2\}$ and $\mathrm{Suff}_{\overline{\alpha}}(w_0) = \{\overline{v_0}, \overline{v_1}\}$. Note that $u_1$ $(v_1)$ must be primitive; otherwise, its primitive root is also an $\alpha$-prefix (resp. $\overline{\alpha}$-suffix) of $w_0$ (Lemma 8) and $w_0$ would not be a $(3,2)$-word any more. As a result, $u_1$ commute with $v_1$ $(u_2)$ if and only if $u_1 = v_1$ (resp. $u_2 = u_1^2$). Recall also that $u_2 \neq v_1$ must hold for $w_0$ to be $(3,2)$-word (Proposition 10). Thus, if $u_2$ and $v_1$ commute, then $u_2 = v_1^2$ and $u_1 = v_1$. In other words, the commutativity between $u_2$ and

$v_1$ is reduced to the commutativity between $u_1$ and $u_2$ and the commutativity between $u_1$ and $v_1$, and hence, is not essential.

Corollary 13 states that $\mathcal{H}_\alpha(w_0) = \{w_0\} \cup \{v_1 w_0, w_0\overline{u_1}, w_0\overline{u_2}\}$. Let us ask the question of whether iterated hairpin completion can generate a same word from $w_0\overline{u_1}$ and $w_0\overline{u_2}$. We partially answer this question in a broader setting for arbitrary $m \geq 3$ and $n \geq 1$.

**Lemma 20.** *Let $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ be a non-crossing $(m,n)$-word for some $m \geq 3$ and $n \geq 1$ with $\mathrm{Pref}_\alpha(w_0) = \{u_0, u_1, \ldots, u_{m-1}\}$. For integers $i, j$ with $1 \leq i < j < m$, if $u_j\alpha \geq_s u_i\alpha$, then $\mathcal{H}_\alpha^*(w_0\overline{u_j}) \subseteq \mathcal{H}_\alpha^*(w_0\overline{u_i})$; otherwise, $\mathcal{H}_\alpha^*(w_0\overline{u_j}) \cap \Sigma^* w_0\overline{u_i}\Sigma^* = \emptyset$.*

*Proof.* The suffix relation means $u_j \in \{u_1, \ldots, u_{j-1}\}u_i$ (Lemma 6) so that let $u_j = xu_i$ for some $x \in \{u_1, \ldots, u_{j-1}\}$. Lemma 12 implies that hairpin completion can extend $w_0\overline{u_i}$ to the right by any of $u_1, \ldots, u_{j-1}$ so that $w_0\overline{u_i} \to_{\mathcal{RH}} w_0\overline{u_i}\ \overline{x} = w_0\overline{u_j}$ is valid. Thus, the inclusion holds. Conversely, if the intersection is not empty, then Theorem 15 implies that $\overline{\alpha}\,\overline{u_j} = \overline{\alpha}\,\overline{u_i}\,\overline{y}$ for some $y \in \Sigma^+$. Due to Lemma 6, this equation gives $y \in \{u_1, \ldots, u_{j-1}\}$; thus, $u_j \in \{u_1, \ldots, u_{j-1}\}u_i$. □

We can employ Lemma 20 in our current setting of non-crossing $(3,2)$-words to observe that if $u_2 = u_1^2$, then $\mathcal{H}_\alpha^*(w_0\overline{u_2}) \subseteq \mathcal{H}_\alpha^*(w_0\overline{u_1})$; otherwise, $\mathcal{H}_\alpha^*(w_0\overline{u_2}) \cap \Sigma^* w_0\overline{u_1}\Sigma^* = \emptyset$. Thus, for example, if $u_2 \neq u_1^2$, then $\mathcal{H}_\alpha^*(w_0\overline{u_2}) \cap \mathcal{H}_\alpha^*(w_0\overline{u_1}) = \emptyset$.

In this subsection, we first prove that the commutativity of $u_1$ with $v_1$ or with $u_2$ is a sufficient condition for $\mathcal{H}_\alpha^*(w_0)$ to be regular.

**Lemma 21.** *If $u_1 = v_1$, then $\mathcal{H}_\alpha^*(w_0)$ is regular.*

*Proof.* Since $u_1 = v_1$, we can let $w_0 = w\overline{u_1}$ with some $w \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$, which is a non-crossing $(3,1)$-word with $u_1, u_2$ being its nonempty $\alpha$-prefix. Lemma 12 implies that $|u_1| + 2|\alpha| < |w|$, which means that hairpin completion can extend $w$ to the right by $\overline{u_1}$ and result in $w_0$. If $|u_2| + 2|\alpha| \leq |w|$, then $w\overline{u_2}$ can be thus generalized, though it is not essential here whether this is possible or not. Let us consider only the case when it is possible. Then $\mathcal{H}_\alpha^*(w)$, which is regular due to Theorem 17, is $\{w\} \cup \mathcal{H}_\alpha^*(w\overline{u_1}) \cup \mathcal{H}_\alpha^*(w\overline{u_2})$. As we have seen above, if $w\overline{u_2} \in \mathcal{H}_\alpha(w)$, then either $\Sigma^* w\overline{u_1}\Sigma^* \cap \mathcal{H}_\alpha^*(w\overline{u_2}) = \emptyset$ or $\mathcal{H}_\alpha^*(w\overline{u_1}) \supseteq \mathcal{H}_\alpha^*(w\overline{u_2})$. In any case, $\mathcal{H}_\alpha^*(w_0) = \mathcal{H}_\alpha^*(w) \cap \Sigma^* w\overline{u_1}\Sigma^*$, and hence, is regular. □

Now it is easy to see that $\mathcal{H}_\alpha^*(w_0)$ is regular when $u_2$ commutes with $v_1$. Since $w_0$ is a $(3,2)$-word, $v_1$ must be primitive and $u_2$ is equal to either $v_1$ or $v_1^2$. In the former case, however, $u_1$ is a proper prefix of $v_1$ so that $w_0 \geq_s \overline{\alpha}\,\overline{u_1}$ and $w_0$ would not be a $(3,2)$-word any more. Thus, the latter must be the case. Then, the prefix $v_1$ of $u_2$, which is the primitive root of $u_2$, is an $\alpha$-prefix of $w_0$ (Lemma 8), and hence, in order for $w_0$ to be a $(3,2)$-word, $u_1 = v_1$ must hold, and this brings the conclusion according to Lemma 21.

**Lemma 22.** *If $u_2 = u_1^2$, then $\mathcal{H}_\alpha^*(w_0)$ is regular.*

*Proof.* Lemma 21 makes it sufficient to consider the case when $u_1$ does not commute with $v_1$. Lemma 12 implies that $\mathcal{H}_\alpha(w_0) = \{w_0, v_1w_0, w_0\overline{u_1}, w_0\overline{u_1}^2\}$ and $w_0\overline{u_1} \rightarrow_{\mathcal{RH}} w_0\overline{u_1}^2$. Hence, $\mathcal{H}^*_\alpha(w_0) = \{w_0\} \cup \mathcal{H}^*_\alpha(v_1w_0) \cup \mathcal{H}^*_\alpha(w_0\overline{u_1})$. We will show the regularity of the second and third terms of this equation.

First, we prove that $\mathcal{H}^*_\alpha(w_0\overline{u_1})$ is regular. One can let $w_0 = u_1w$, where $w \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ is a $(2,2)$-word with $\mathrm{Pref}_\alpha(w) = \{\lambda, u_1\}$ and $\mathrm{Suff}_{\overline{\alpha}}(w) = \{\lambda, \overline{v_1}\}$. It is left to the readers to check that

$$\mathcal{H}^*_\alpha(w) = w\overline{u_1}^* \cup \mathcal{H}^*_\alpha(u_1w\overline{u_1}) \cup \mathcal{H}^*_\alpha(u_1v_1w\overline{u_1}) \cup \mathcal{H}^*_\alpha(v_1w).$$

As done in the proof of Lemma 21, the non-commutativity of $u_1$ with $v_1$ implies that $(\mathcal{H}^*_\alpha(u_1v_1w\overline{u_1}) \cup \mathcal{H}^*_\alpha(v_1w)) \cap \Sigma^*u_1w\Sigma^* = \emptyset$. Thus, $\mathcal{H}^*_\alpha(w) \cap \Sigma^*u_1w\Sigma^* = \mathcal{H}^*_\alpha(w_0\overline{u_1})$. Since $w$ is a non-crossing $(2,2)$-word, $\mathcal{H}^*_\alpha(w)$ is regular (Theorem 19), and hence, $\mathcal{H}^*_\alpha(w_0\overline{u_1})$ is regular, too.

Next, we prove the regularity of $\mathcal{H}^*_\alpha(v_1w_0)$. We can let $w_0 = w'\overline{v_1}$ for some $(3,1)$-word $w'$. This means that $v_1w'$ is a $(4,1)$-word with $\mathrm{Pref}_\alpha(v_1w') = \{\lambda, v_1, v_1u_1, v_1u_1^2\}$ and the empty $\overline{\alpha}$-suffix. Thus,

$$\mathcal{H}^*_\alpha(v_1w') = \{v_1w'\} \cup \mathcal{H}^*_\alpha(v_1w'\overline{v_1}) \cup \mathcal{H}^*_\alpha(v_1w'\overline{u_1}\,\overline{v_1}) \cup \mathcal{H}^*_\alpha(v_1w'\overline{u_1}^2\overline{v_1}).$$

Using the essentially same argument as above, we obtain $\mathcal{H}^*_\alpha(v_1w') \cap \Sigma^*v_1w'\overline{v_1}\Sigma^* = \mathcal{H}^*_\alpha(v_1w_0)$. Since the iterated hairpin completion of non-crossing $(4,1)$-word is regular (Theorem 17), $\mathcal{H}^*_\alpha(v_1w')$ is regular and this implies the regularity of $\mathcal{H}^*_\alpha(v_1w_0)$. □

Thus, any two of $u_1, u_2, v_1$ must not commute in order for $\mathcal{H}^*_\alpha(w_0)$ not to be regular. Let us add one more sufficient condition for $\mathcal{H}^*_\alpha(w_0)$ to be regular.

**Lemma 23.** *If $u_2 = u_1v_1$, then $\mathcal{H}^*_\alpha(w_0)$ is regular.*

*Proof.* Due to Lemma 21, it suffices to prove this lemma under the assumption $u_1 \neq v_1$, which is equivalent to that $u_1$ does not commute with $v_1$ under our problem setting.

Note that $\mathcal{H}^*_\alpha(w_0) = \{w_0\} \cup \mathcal{H}^*_\alpha(v_1w_0) \cup \mathcal{H}^*_\alpha(w_0\overline{u_1}) \cup \mathcal{H}^*_\alpha(w_0\overline{v_1}\,\overline{u_1})$. As done before, we will check that the second, third, and fourth terms of the union above are regular. The regularity of the third one follows from $\mathrm{Pref}_\alpha(w_0\overline{u_1}) = \{\lambda, u_1, u_1v_1\}$, $\mathrm{Suff}_{\overline{\alpha}}(w_0\overline{u_1}) = \{\lambda, \overline{u_1}, \overline{v_1}\,\overline{u_1}\}$, and Corollary 18.

In order to check that the second term is regular, let $w_0 = w\overline{v_1}$, where $w$ is a non-crossing $(3,1)$-word. Then $v_1w$ is a $(4,1)$-word, and

$$\mathcal{H}^*_\alpha(v_1w) = \{v_1w\} \cup \mathcal{H}^*_\alpha(v_1w\overline{v_1}) \cup \mathcal{H}^*_\alpha(v_1w\overline{u_1}\,\overline{v_1}) \cup \mathcal{H}^*_\alpha(v_1w\overline{v_1}\,\overline{u_1}\,\overline{v_1}).$$

Since $v_1w\overline{v_1} \rightarrow_{\mathcal{RH}} v_1w\overline{v_1}\,\overline{u_1}\,\overline{v_1}$ and $\mathcal{H}^*_\alpha(v_1w\overline{u_1}\,\overline{v_1}) \cap \Sigma^*v_1w\overline{v_1}\Sigma^* = \emptyset$, we have $\mathcal{H}^*_\alpha(v_1w_0) = \mathcal{H}_\alpha(v_1w\overline{v_1}) = \mathcal{H}^*_\alpha(v_1w) \cap \Sigma^*v_1w\overline{v_1}\Sigma^*$. The regularity of $\mathcal{H}^*_\alpha(v_1w)$ is due to Theorem 17 so that $\mathcal{H}_\alpha(v_1w_0)$ is regular.

What remains to be considered is the fourth term. One can let $w_0\overline{v_1}\,\overline{u_1} = u_1v_1w'$ for some non-crossing $(1,4)$-word $w'$. Then $\mathcal{H}^*_\alpha(w') = \{w'\} \cup \mathcal{H}^*_\alpha(u_1w') \cup \mathcal{H}^*_\alpha(u_1v_1w') \cup \mathcal{H}^*_\alpha(u_1v_1^2w')$, and we can easily see that $\mathcal{H}^*_\alpha(u_1v_1w') = \mathcal{H}^*_\alpha(w') \cap \Sigma^*u_1v_1w'\Sigma^*$. Since $w'$ is a $(1,4)$-word, $\mathcal{H}^*_\alpha(w')$ is regular, and hence, $\mathcal{H}^*_\alpha(w_0\overline{v_1}\,\overline{u_1})$ is regular, too. □

**Theorem 24.** *Let $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ be a non-crossing $(3,2)$-word with $\mathrm{Pref}_\alpha(w_0) = \{\lambda, u_1, u_2\}$ and $\mathrm{Suff}_{\overline{\alpha}}(w_0) = \{\lambda, \overline{v_1}\}$. Then $\mathcal{H}_\alpha^*(w_0)$ is regular if and only if one of the following three conditions holds: (1) $u_1$ commutes with $v_1$; (2) $u_1$ commutes with $u_2$; or (3) $u_2 = u_1 v_1$.*

*Proof.* Let $R = \{u_2 u_1^i v_1 w_0 \overline{u_1}^j \overline{u_2} \mid i, j \geq 2\}$, which is a regular language. We will prove that, under the assumption that none of the conditions 1-3 holds, $\mathcal{H}_\alpha^*(w_0) \cap R = \{u_2 u_1^i v_1 w_0 \overline{u_1}^i \overline{u_2} \mid i \geq 2\}$ holds so that $\mathcal{H}_\alpha^*(w_0)$ is not regular. Let us denote this intersection by $L$.

As mentioned previously, if the second condition does not hold (this is equivalent to $u_2 \neq u_1^2$), then $\mathcal{H}_\alpha^*(w_0 \overline{u_2})$ cannot contain any word in the above intersection. Thus, $L = (\mathcal{H}_\alpha^*(w_0 \overline{u_1}) \cap R) \cup (\mathcal{H}_\alpha^*(v_1 w_0) \cap R)$. Using Lemmas 6 and 7, we can easily prove the emptiness of the second intersection of the above sum. This check is left to the reader, and the authors recommend them to check at least $\mathcal{H}_\alpha^*(v_1 w_0 \overline{u_2}\, \overline{v_1}) \cap R = \emptyset$ because this check involves the important fact that $\overline{\alpha}\,\overline{u_1} \leq_p \overline{\alpha}\,\overline{u_2}$ implies $u_2 = u_1^2$ (Lemma 6) and causes a contradiction. As a result, we have $L = \mathcal{H}_\alpha^*(w_0 \overline{u_1}) \cap R$. Informally speaking, in order to produce a word in $R$ from $w_0$, we first have to extend $w_0$ to the right by $\overline{u_1}$.

Now we can extend $w_0 \overline{u_1}$ to the right by $\overline{u_1}$ $i$-times to obtain $w_0 \overline{u_1}^i$. If this obtained word is extended to the left, then the word will be in $u_1\{u_1, u_2, v_1\}^* w_0 \{u_1, u_2, v_1\}^* \overline{u_1}$. Note that any word obtained from this word by iterated hairpin completion is also in $u_1\{u_1, u_2, v_1\}^* w_0 \{u_1, u_2, v_1\}^* \overline{u_1}$. Let us claim that $u_1\{u_1, u_2, v_1\}^* w_0 \{u_1, u_2, v_1\}^* \overline{u_1} \cap u_2 \Sigma^* w_0 \Sigma^* \overline{u_2} = \emptyset$. Indeed, if the intersection were not empty, then $u_2 \alpha \leq_p u_1 x \alpha$ for some $x \in \{u_1, u_2, v_1\}^+$. Due to Lemma 5, $u_2 \in u_1\{u_1, v_1\}^+$, but actually we can say $u_2 \in u_1\{u_1, v_1\}$ for $u_2$ is the second shortest nonempty $\alpha$-prefix of $w_0$. However, this means that either the condition 1 or 2 holds, and contradicts our assumption. Thus, we have only one choice; extending $w_0 \overline{u_1}^i$ to the right by $\overline{u_2}$.

As mentioned above, $\overline{\alpha}\,\overline{u_1} \leq_p \overline{\alpha}\,\overline{u_2}$ cannot hold so that we cannot extend $w_0 \overline{u_1}^i \overline{u_2}$ further to the right to obtain a word in $R$. Thus, we should extend this word to the left either by $u_2 u_1^j$ for some $j \leq i$ or by $u_2 u_1^i v_1$. Lemmas 6 and 7 prove that the former choice will not lead us to any word in $R$. Now it suffices to mention that extending $u_2 u_1^i v_1 w_0 \overline{u_1}^i \overline{u_2}$ further to the left will not produce any word in $R$ because such an extension force the contradictory relation $\overline{\alpha}\,\overline{u_1} \leq_p \overline{\alpha}\,\overline{u_2}$ to hold. □

# 5 Conclusion

In this paper, we focused on finding conditions that a word $w_0 \in \alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$ must satisfy so that its iterated hairpin completion $\mathcal{H}_\alpha^*(w_0)$ is a regular language. We classified the set of all non-crossing words according to the number $m$ of occurrences of $\alpha$ and the number $n$ of occurrences of $\overline{\alpha}$ on a given word. For the cases when $n = 1$ and when $m = n = 2$, we proved that the iterated hairpin completion of a non-crossing $(m, n)$-word is regular. We also found a necessary and sufficient condition under which the iterated hairpin completion of a non-crossing $(3,2)$-word is regular. This approach can be generalized to arbitrary non-crossing $(m, n)$-words, with the cases $(m, 1)$ and $(2, 2)$ being the induction base of an inductive proof. Future works include considering the same problem for crossing-words. In this case, Lemma 12 or Theorem 15 does not hold any more, and hence, it may get harder to analyze the derivation processes of how a word is obtained from a given word $w_0$ by iterated hairpin completion. In addition, we investigated only the case when the suffix of length $k$ of an initial word $w_0$ is the complement of its prefix of the same length, but we eventually have to consider $w_0$ in $\alpha\Sigma^* \cap \Sigma^*\overline{\beta}$, where $\beta$ might not be equal to $\alpha$ (double-primer hairpin completion). We can easily observe that one-step hairpin completion with respect to $\alpha$ ($\beta$) derives a word in $\beta\Sigma^* \cap \Sigma^*\overline{\beta}$ (resp. $\alpha\Sigma^* \cap \Sigma^*\overline{\alpha}$) from $w_0$. Thus, results obtained in this study of single-primer hairpin completion are an important step towards this most general setting of the regularity test problem of iterated hairpin completion of a single word. Another direction of research is to consider stopper sequences as in Whiplash PCR [7, 18].

## References


[1] Adleman, L. M.: Molecular Computation of Solutions to Combinatorial Problems, *Science*, **266**(5187), 1994, 1021–1024.

[2] Arita, M., Kobayashi, S.: DNA Sequence Design Using Templates, *New Generation Computing*, **20**, 2002, 263–277.

[3] Cheptea, D., Martín-Vide, C., Mitrana, V.: A New Operation on Words Suggested by DNA Biochemistry: Hairpin Completion, *Transgressive Computing* (J.-G. Dumas, Ed.), 2006, 216-228.

[4] Choffrut, C., Karhumäki, J.: Combinatorics of Words, in: *Handbook of Formal Languages* (G. Rozenberg, A. Salomaa, Eds.), vol. 1, Springer-Verlag, Berlin-Heidelberg-New York, 1997, 329–438.

[5] Diekert, V., Kopecki, S.: Complexity Results and the Growths of Hairpin Completions of Regular Languages (Extended Abstract), *Proc. of CIAA 2010* (M. Domaratzki, K. Salomaa, Eds.), LNCS 6482, Springer, 2011, 105-114.

[6] Fine, N. J., Wilf, H. S.: Uniqueness Theorem for Periodic Functions, *Proceedings of the American Mathematical Society*, **16**(1), 1965, 109–114.

[7] Hagiya, M., Arita, M., Kiga, D., Sakamoto, K., Yokoyama, S.: Towards Parallel Evaluation and Learning of Boolean $\mu$-Formulas with Molecules, *Proc. of 3rd DIMACS Workshop on DNA Based Computers* (H. Rubin, D. Wood, Eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science 48, 1999, 57-72.

[8] Ito, M., Leupold, P., Manea, F., Mitrana, V.: Bounded Hairpin Completion, *Information and Computation*, **209**(3), 2011, 471–485.

[9] Jonoska, N., Kephart, D., Mahalingam, K.: Generating DNA Codewords, *Congressus Numerantium*, **156**, 2002, 99–110.

[10] Jonoska, N., Mahalingam, K.: Languages of DNA Based Code Words, *Proc. of DNA Computing 9* (J. Chen, J. Reif, Eds.), LNCS 2943, Springer, 2004, 61-73.

[11] Kameda, A., Yamamoto, M., Ohuchi, A., Yaegashi, S., Hagiya, M.: Unravel Four Hairpins!, *Natural Computing*, **7**, 2008, 287–298.

[12] Kopecki, S.: On Iterated Hairpin Completion, *Theoretical Computer Science*, **412**(29), 2011, 3629–3638.

[13] Manea, F., Martín-Vide, C., Mitrana, V.: On Some Algorithmic Problems Regarding the Hairpin Completion, *Discrete Applied Mathematics*, **157**, 2009, 2143–2152.

[14] Manea, F., Mitrana, V., Yokomori, T.: Two Complementary Operations Inspired by the DNA Hairpin Formations: Completion and Reduction, *Theoretical Computer Science*, **410**(4-5), 2009, 417–425.

[15] Manea, F., Mitrana, V., Yokomori, T.: Some Remarks on the Hairpin Completion, *International Journal of Foundations of Computer Science*, **21**(5), 2010, 859–872.

[16] Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., Sugino, A.: Mechanism of DNA Chain Growth, I. Possible Discontinuity and Unusual Secondary Structure of Newly Synthesized Chains, *Proceedings of the National Academy of Sciences of the United States of America*, **59**(2), 1968, 598–605.

[17] Păun, G., Rozenberg, G., Yokomori, T.: Hairpin Languages, *International Journal of Foundations of Computer Science*, **12**(6), 2001, 837–847.

[18] Sakamoto, K., Kiga, D., Komiya, K., Gouzu, H., Yokoyama, S., Ikeda, S., Sugiyama, H., Hagiya, M.: State Transitions by Molecules, *Biosystems*, **52**(1-3), 1999, 81–91.

[19] Takinoue, M., Suyama, A.: Molecular Reactions for a Molecular Memory Based on Hairpin DNA, *Chem-Bio Informatics Journal*, **4**, 2004, 93–100.

[20] Takinoue, M., Suyama, A.: Hairpin-DNA Memory Using Molecular Addressing, *Small*, **2**(11), 2006, 1244–1247.

[21] Wilson, K. S., von Hippel, P. H.: Transcription Termination at Intrinsic Terminators: The Role of the RNA Hairpin, *Proceedings of the National Academy of Sciences of the United States of America*, **92**(19), 1995, 8793–8797.