Communications in Combinatorics and Optimization
Vol. 9, No. 2 (2024), pp. 263-277
https://doi.org/10.22049/cco.2022.27917.1395

CCO
Commun. Comb. Optim.

Research Article

The length of the longest sequence of consecutive FS-double squares in a word

Maithilee Patawar[1,*] and Kalpesh Kapoor[2]

[1]Department of Computer Science & Engineering, Indian Institute of Technology Guwahati, India
maith176101104@iitg.ac.in

[2]Department of Mathematics, Indian Institute of Technology Guwahati, India
kalpesh@iitg.ac.in

Received: 10 July 2022; Accepted: 11 December 2022
Published Online: 25 December 2022

Abstract: A square is a concatenation of two identical words, and a word w is said to have a square yy if w can be written as $xyyz$ for some words x and z. It is known that the ratio of the number of distinct squares in a word to its length is less than two, and any location of a word could begin with two distinct squares which are appearing in the word for the last time. A square whose first location starts with the last occurrence of two distinct squares is an FS-double square. We explore and identify the conditions under which a sequence of locations in a word starts with FS-double squares. We first find the structure of a word that begins with two consecutive FS-double squares and obtain its properties that enable us to extend the sequence of FS-double squares. It is proved that the length of the longest sequence of consecutive FS-double squares in a word of length n is at most $\frac{n}{7}$. We show that the squares in the longest sequence of consecutive FS-double squares are conjugates.

Keywords: Distinct squares, FS-double squares, Repetitions, Word combinatorics

AMS Subject classification: 68R15

1. Introduction

A word is a finite sequence of symbols or letters drawn from a nonempty finite set. A repetition in a word is of the form u^m for some nonempty word u and an integer $m > 1$. The smallest repetition obtained with $m = 2$ is known as a square and is a concatenation of two identical words. The study of squares reveals interesting

* *Corresponding Author*

properties of words, and is a well-researched topic in word combinatorics, see for example [2, 5, 15].

Fraenkel and Simpson conjectured that the number of distinct squares in a word is always less than its length [11]. They considered the last occurrence of a square and showed that a location in a word could begin with the last occurrences of at most two distinct squares. An alternate proof for the same result is discussed in [13]. We refer to the conjecture for the bound on the maximum number of distinct squares in a word as the "square conjecture".

An obvious direction to precisely compute the number of distinct squares in a word is to find the maximum number of possible locations starting with the last instances of two distinct squares. In [14], efforts are made in this direction which achieved a new upper bound $2n - \Theta(\log n)$ for the square conjecture. The bound is obtained by studying the words containing "runs of $2'$s", where a '2' indicates a location that starts with two distinct squares such that these squares never repeat after the location. Given such a location, Deza et al. [9] refer to the longer square starting at the location as an FS-double square. For such a square that starts at a location, say i, any FS-double square starting after the location i is divided into five categories. The characteristics of the words belonging to these categories are studied to obtain an upper bound on the number of FS-double squares. It is found that a word, w, contains at most $\lfloor\frac{5|w|}{6}\rfloor$ FS-double squares and $\lfloor\frac{11|w|}{6}\rfloor$ distinct squares [9]. A recent work [18], published in arXiv, redefines the classification of a pair of FS-double squares and has proved an improved upper bound of $\frac{3|w|}{2}$ for the square conjecture. The latest work [4], again published in arXiv, validated the square conjecture using a graph based approach.

A d-step conjecture [7], predicts that the number of distinct primitively-rooted squares in a word, w, is less than $|w| - d$, where d is the number of distinct letters in w. The conjecture is still open, but it is supported by a computational framework that attempts to produce words containing a large number of distinct primitively-rooted squares [8]. On the contrary, a framework with the objective of disproving the conjecture by finding a counterexample is presented in [12].

The ratio of the number of distinct squares in a word to its length is referred to as square density, and the relation between the square densities of words over a varying alphabet is explored in [16]. This work also showed that the result on binary words is sufficient to solve the square conjecture or the d-step conjecture. In [3], Blanchet-Sadri et al. designed a family of words that attempts to maximize the number of FS-double squares. The FS-double squares in these words start at consecutive locations and are of the same size. It is also shown that a word containing m such equal length FS-double squares has at least $7m + 3$ letters. Further, the density of such a word is at most $\frac{5}{6}$.

The earlier works indicate that having a large number of squares starting at consecutive locations can possibly increase the square density of a word. The best-known words that have been constructed to study the square conjecture have such sequences [11, 17]. In these words, the squares at consecutive locations are conjugates obtained by extending a square by its prefix. It is possible to extend an FS-double square in

a similar fashion to introduce consecutive FS-double squares. Clearly, the bounds for the square conjecture are mainly dominated by two factors; first, a structure of words in which many squares appear at consecutive locations, and, secondly, a structure of words having as many FS-double squares as possible.
However, based on the discussion in [13], we suspect that the presence of consecutive FS-double squares in a word reduces its square density. This motivates us to explore the words with a sequence of locations starting with FS-double squares. We show that such words are possible only under specific conditions. Our goal in this study is to identify the ways that maximize consecutive FS-double squares. In particular, the following are the main contributions of our work.

(a) Identify structures of words having FS-double squares at consecutive locations.

(b) We give an upper bound for the maximum number of consecutive FS-double squares in a word.

The rest of the paper is organized as follows. In the next section, we introduce basic definitions and give an overview of prior work. In Section 3, we consider all possible conditions for two consecutive FS-double squares. The feasibility of these conditions is verified against the properties of FS-double squares, and it is shown that only two structures are feasible. Further, in Section 4 we explore the properties of two structures discovered in the previous section and employ them to obtain the length of the longest sequence of FS-double squares. Finally, we summarize in Section 5.

2. Definitions and Background

A word is a concatenation of finite number of symbols drawn from an alphabet Σ. The concatenation of two words w_1 and w_2 is represented as $w_1 \cdot w_2$ or simply $w_1 w_2$. The '$\cdot$' operation is associative. The number of letters in a word w is denoted by $|w|$. The word with length zero is denoted by ϵ. The i^{th} power of a word w, denoted by w^i, is defined recursively as $w^0 = \epsilon$ and $w^i = w^{i-1} \cdot w$ for $i \geq 1$.
Define Σ^* (Σ^+) as set of all (nonempty) words over Σ. If $w = pqr$ then the words p, q and r are called prefix, factor and suffix, respectively, of w. A proper prefix (factor, suffix) of a word, w, is a prefix (factor, suffix) which is not equal to ϵ or w. The longest common prefix of two words w_1 and w_2 is denoted as $lcp(w_1, w_2)$. A word $\widetilde{w}$ is said to be a conjugate of a word w if and only if $w = uv$ and $\widetilde{w} = vu$ for some $u, v \in \Sigma^+$. Note that a conjugate of a square is also a square.
A nonempty word, w, is primitive if and only if whenever $w = u^k$ implies $k = 1$. If a word, w, satisfies the relation $w = u^k$ for some word u and an integer $k \geq 2$, then it is non-primitive. A square is a word of the form uu where the nonempty word u is referred to as its root. A square whose root is primitive is called as a primitively-rooted square. The following lemma is a consequence of Lemma 1.4 of [6].

Lemma 1. *If x is primitive, then no two conjugates of x are the same.*

Let $w = a_1a_2 \dots a_n$ be a word. A square, u^2, whose last occurrence in a word begins at a location i is denoted as u_i^2. Define s_i to be the number of those squares that start at location i in w which do not start at location j where $i < j \leq n$. It is known that for any word the value of s_i is 0, 1 or 2 for $1 \leq i \leq n$. As a corollary, the number of distinct squares in a word of length n is bounded by $2n$ for a word defined over any finite alphabet [11].
Deza et al. [9] studied the properties of words with $s_i = 2$, that is, words in which two distinct squares begin at location i and these are appearing in the word for the last time. They referred to the longer square that starts at a location i with $s_i = 2$ as an "FS-double Square". For such a square that begins at location i, we use sq_i^2 and SQ_i^2 to denote the shorter and the longer squares, respectively. In the same work, Lemma 6 defines the term balanced double square along its structure which is further extended after Definition 7 to get the following structure of FS-double squares.

Lemma 2 ([9]). *The roots, sq_i and SQ_i, of an FS-double square starting at location i have the following structure:*

$$sq_i = (xy)^{p+q}(x) \qquad\qquad SQ_i = (xy)^{p+q}(x)(xy)^p \tag{1}$$

where p and q are integers such that $p \geq 1, q \geq 0$, $x, y \neq \epsilon$ and xy is a primitive word.

The same work that introduced the above lemma also defined the concept of "mate" of an FS-double square in a word. The next section describes the classification of mates based on the lengths and locations of FS-double squares.

2.1. Mates of an FS-double Square

For an FS-double square SQ_1^2, Deza et al. [9] categorized another FS-double square $SQ_k^2, k > 1$ into five types based on the value of k and the sizes of roots $\{sq_1, SQ_1\}, \{sq_k, SQ_k\}$. For $k < (p+q-1)|xy| + |lcp(xy, yx)|$, an FS-double square SQ_k^2 is one of the following mates of SQ_1^2 if it satisfies the necessary conditions.

(a) α-mate: The root SQ_k is a conjugate of SQ_1 which gives $|SQ_1| = |SQ_k|$. The condition $|sq_1| = |sq_k|$ is further added in [18].

(b) β-mate: The FS-double squares satisfy the relation $|SQ_1| = |SQ_k|$ and the root $sq_k = (\widetilde{xy})^i \widetilde{x}$ for some integer $i \in [2, p+q-1]$. The words $\widetilde{xy}$ and $\widetilde{x}$ are conjugates of xy and x, respectively. It implies that $|sq_1| > |sq_k|$.

(c) γ-mate: Here, $k < p + q|xy|$ and $|sq_k| = |SQ_1|$.

(d) δ-mate: The lengths of the roots satisfy $|sq_k| > |SQ_1|$.

An ϵ-mate of SQ_1^2 starts after $(p+q-1)|xy| + |lcp(xy, yx)|$ locations. Since we are interested in consecutive FS-double squares, we do not need to consider ϵ-mate in this paper. In addition to these five mates, Thierry in [18] described the other two mates, named η and ζ-mates. However, an η-mate or a ζ-mate of an FS-double square SQ_k^2 can never start at location $k+1$. So, the two FS-double squares beginning at adjacent locations are only of the above four kinds. In the next section, we show that these FS-double squares are either α or δ-mates and cannot be β or γ-mates.

3. Structure of 2FS Squares

We first study the structure of the word having the smallest possible sequence of consecutive FS-double squares. A word with the smallest such sequence begins with two FS-double squares where these FS-double squares start at consecutive locations. We call it a 2FS square. We understand various properties of a 2FS square using the following results on FS-double squares and primitively-rooted squares.

Lemma 3 (Two Squares Factorization Lemma [1]). *Let an FS-double square, SQ_i^2, begins with a shorter square, sq_i^2, such that $sq_i = (xy)^{p+q}x$ and $SQ_i = sq_i(xy)^p$ where $x, y \in \Sigma^+$ and the integers p, q satisfy $p \geq 1, q \geq 0$ then,*

(a) SQ_i is primitively-rooted, and

(b) sq_i is primitively-rooted if $p+q > 1$.

We now see a lemma that gives a way to introduce distinct squares of equal lengths that start at consecutive locations by extending a primitively-rooted square. The obtained squares, in this case, are conjugates of each other. The lemma is a consequence of a well-known result named Periodicity Lemma [10].

Lemma 4 ([17]). *Let $w = uu$ be a primitively-rooted square such that $|u| > 1$ and v be a proper prefix of u. Then, the suffix v in the word wv introduces $|v|$ conjugates of uu.*

It is also possible to extend a square where the root of the square is non-primitive. In such cases, the newly introduced squares may not always be distinct.

Lemma 5. *Let uu and vv be two squares with $|u| = |v|$ that appear at consecutive locations in a word. Then, $u = \widetilde{v}$ and $uu = \widetilde{vv}$. Further, $ua = av$ and $uua = avv$ for some $a \in \Sigma$.*

Proof. Let uu and vv be two consecutive squares in a word where the former square ends first. Assume u begins with a letter a and v ends with a letter b such that $u = au', v = v'b$. The squares uu and vv appear at consecutive locations leading to a relation $uub = avv$. So, the highlighted factors in $a\mathbf{u'}\mathbf{a}u'b = uub$ and $a\mathbf{v'}\mathbf{b}v'b = avv$ must be equal. Therefore, $u' = v'$, $a = b$, $uu = av'av'$ and $vv = v'av'a$. Thus, the squares uu and vv are conjugates. □

In Subsection 2.1, we discussed the definitions of different mates that are given in [9]. Accordingly, an α-mate of an FS-double square SQ_1^2 refers to an FS-double square

SQ_k^2 where $k > 1$ and $|sq_1| = |sq_k|, |SQ_1| = |SQ_k|$. In the next subsection, we show that for $k = 2$, $|SQ_1| = |SQ_2| \iff |sq_1| = |sq_2|$.

3.1. Equal 2FS Squares

Let $w = a_1 \ldots a_n$ be a word. We refer to a factor starting at location i with $s_i = s_{i+1} = 2$ as 2FS square if the FS-double square starting at location $i + 1$ is a suffix of w. Consider a 2FS square in which $s_1 = s_2 = 2$ with (sq_1, SQ_1) and (sq_2, SQ_2) being the two respective pairs of the roots of FS-double squares. The 2FS square is an equal 2FS square if it satisfies the relation $|sq_1| = |sq_2|$ and $|SQ_1| = |SQ_2|$. Otherwise, we call it an unequal 2FS square. According to the structure of an FS-double square provided in Lemma 2, any 2FS square satisfies the relation $|sq_1| < |SQ_1|$ and $|sq_2| < |SQ_2|$. In this section, we further compare the lengths of the roots sq_1, sq_2, SQ_1 and SQ_2 in detail. Now, assume the FS-double squares SQ_1^2 and SQ_2^2 have the following structures for the given 2FS square.

$$\begin{aligned} SQ_1 &= (xy)^{p+q}(x)(xy)^p \\ SQ_2 &= (uv)^{p'+q'}(u)(uv)^{p'} \end{aligned}$$

In the rest of the paper, we assume that SQ_1 begins with a letter a such that $x = ax'$.

Lemma 6. *Let w be a 2FS square with $|SQ_1| = |SQ_2|$. Then, $|sq_1| = |sq_2|$.*

Proof. Let SQ_1^2 begins with a letter a such that $x = ax'$ for some word x'. We have SQ_2^2 as follows where $x'ya$ is primitive.

$$SQ_2^2 = (x'ya)^{p_1}(x'a)(x'ya)^{p_2} \cdot (x'ya)^{p_1}(x'a)(x'ya)^{p_2} \tag{2}$$

If $|sq_2| < |sq_1|$, then $|sq_1| \neq p_1|x'ya|$. Otherwise, for some suffix s of $x'ya$ we get the following relation where s is one of the prefixes of $x'ya$.

$$sq_2 : (x'ya)^{p_1} = x'a(x'ya)^{p_2}(x'ya)^{p_1-p_2-1}s \tag{3}$$

In the above equation, if s is a proper prefix of $x'ya$ then equating the $|x'ya|$ length suffixes from both side implies that two conjugates of $x'ya$ are the same. This contradicts Lemma 1. If $|s| = 0$, then the prefixes of equation (3) violates Lemma 1 unless $x'ya = x'a$. However, $x'ya = x'a$ implies $|y| = 0$ which is not acceptable. Now, we can obtain the next relation with the constraint $|sq_2| < p_1|x'ya|$.

$$(x'ya)^{p_1-k}s_1 = s_2(x'ya)^{k-1}(x'a)(x'ya)^{p_3}s_3 \tag{4}$$

In order to adhere Lemma 1, $s_1 = s_3$. Moreover, $s_2 = x'ya$ if $k - 1 > 0$. This constraint gives $(x'ya)^{p_1-k} = (x'ya)^k(x'a)(x'ya)^{p_3}$ which inevitably requires to satisfy the condition $x'ya = x'a$. This gives a contradiction. So, the case $k - 1 = 0$ requires to satisfy equation (4). We substitute $k = 1$ in the said equation to get following relation.

$$(x'ya)^{p_1-1} \cancel{s_1} = s_2 x'a(x'ya)^{p_3} \cancel{s_3} \tag{5}$$

Here, the words s_1, s_2 and s_3 are non-empty words such that $s_1 s_2 = x'ya$ (see the structure of SQ_2^2 in equation (2)) and s_1, s_3 are the same to avoid overlapping of conjugates of $x'ya$. While a non-empty s_1 violates Lemma 1, the relation $|s_1| = 0$ again implies $|y| = 0$ contradicting the given assumption. Hence, $|sq_2| \geq |sq_1|$.
The structure of SQ_2^2 when $|sq_2| > |sq_1|$ is shown in Fig. 1. The dotted line in the figure represents the last occurrence of sq_2 in sq_2^2. The root sq_2 shown with a thick line in the figure must has a suffix $x'ya$ (to satisfy Lemma 1).

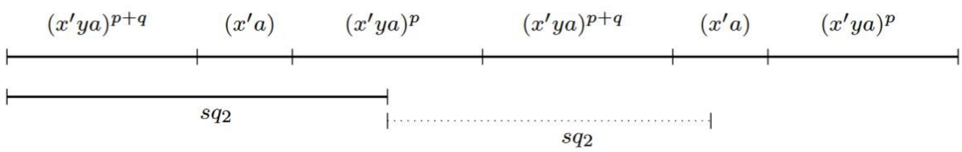

Figure 1. **FS-double square SQ_2^2 when $|sq_2| > |sq_1|$**

So, the first occurrence of the root sq_2 in sq_2^2 ends with $(x'ya)^{p_1}$ for some integer p_1 such that $1 \leq p_1 < p$. The next equations gives the possible structures of the root sq_2 where the LHS and the RHS of each equation indicate the first and the last occurrence, respectively, of sq_2. Assume $r < 2p + q - p_1$ and $p_2 < p$.

$$(x'ya)^{p+q}(x'a)(x'ya)^{p_1} = (x'ya)^{p+q+r} \tag{6}$$

$$(x'ya)^{p+q}(x'a)(x'ya)^{p_1} = (x'ya)^{2p+q-p_1}(x'a)(x'ya)^{p_2} \tag{7}$$

The second occurrence of sq_2 marked with dotted lines in Fig. 1 cannot end somewhere in the factor $x'a$. This is because the two occurrences of sq_2 end with different conjugates of $x'ya$. To avoid the overlap of $x'ya$ with any of its conjugate, equations (6) and (7) need to satisfy the relation $x'a = x'ya$ implying $|y| = 0$. This in unacceptable, thus the FS-double square SQ_2^2 never satisfies the relation $|sq_2| > |sq_1|$. Thus, the only possible relation between lengths of the roots sq_1 and sq_2 is $|sq_1| = |sq_2|$. Applying Lemma 5, sq_1^2 and sq_2^2 are conjugates. □

Lemma 7. *Let w be a 2FS square where $|SQ_1| = |SQ_2|$. Then, SQ_1^2 and SQ_2^2 are α-mates.*

Proof. Since $|SQ_1| = |SQ_2|$, so SQ_1^2 and SQ_2^2 can be α or β-mates. A β-mate is possible in case $|sq_1| > |sq_2|$. From Lemma 6, the roots of the shorter squares only satisfy the relation $|sq_1| = |sq_2|$. So, SQ_1^2 and SQ_2^2 are α-mates. □

Theorem 1 (Equal 2FS Squares). *In a 2FS square,* $|SQ_1| = |SQ_2|$ *if and only if* $|sq_1| = |sq_2|$.

Proof. This follows from Lemma 6 and 7. □

Corollary 1. *Let* w *be a 2FS square. Then,* SQ_1^2 *and* SQ_2^2 *are conjugates if and only if* sq_1^2 *and* sq_2^2 *are conjugates.*

Proof. The statement follows from Lemma 5 and Theorem 1. □

We now explore the condition that is needed to have an equal 2FS square. Suppose w is an FS-double square with sq_1^2 and SQ_1^2. It is possible to extend w to obtain a 2FS square. According to Lemmas 5 and 7, the shorter squares among the two consecutive FS-double squares in an equal 2FS square are conjugates. Therefore, the word w must have a square at the second location, which is a conjugate of sq_1^2.

$$\begin{aligned} SQ_1^2 &= (ax'y)^{p+q} \cdot (ax') \cdot (ax'y)^p \cdot (ax'y)^{p+q} \cdot (ax') \cdot (ax'y)^p \\ &= \underline{(ax'y)^{p+q} \cdot (ax') \cdot (ax'y)^{p+q} \cdot (ax'}y)^p \cdot (ax') \cdot (ax'y)^p \\ &= a \cdot \underbrace{(x'ya)^{p+q} \cdot (x'a) \cdot (x'ya)^{p+q} \cdot x'(y}_{sq_2^2 \implies y \text{ begins with } a} ax')^p \cdot (ax'y)^p \end{aligned} \tag{8}$$

The result of Theorem 1 can also be deduced from Lemma 2 given in [14] (see Lemma 8 below). An obvious question would be why to go into the details of the structures of consecutive FS-double squares while the said lemma discards the possibility of starting a β-mate adjacent to a location starting with the FS-double square. The reason is that we want to highlight the relationship between consecutive FS-double squares and the words x, y shown in equation (8).

Lemma 8 ([14]). *Assume a word* w *begins with two rightmost squares* SQ_1^2 *and* sq_1^2 *such that* $|SQ_1| > |sq_1|$. *If* u^2 *is the rightmost square beginning from the second location of* w, *then* $|u| \in \{|sq_1|, |SQ_2|\}$ *or* $|u| \geq 2|SQ_1|$.

As can be observed from the equation (8), the necessary condition to have an equal 2FS square is that $lcp(x, y)$ must be nonempty. In the following lemma, we count the maximum number of consecutive FS-double squares that are conjugates.

Lemma 9. *Let a word* w *that begins with* i *consecutive FS-double squares such that* $|SQ_1| = |SQ_2| = \cdots = |SQ_i|$ *for some positive integer* i. *Then, the next relation holds.*

$$i \leq \begin{cases} |lcp(xy, yx)| + 1 & \text{if } q > 0 \\ min(|lcp(xy, yx)| + 1, |x|) & \text{otherwise} \end{cases}$$

Proof. As the given consecutive FS-double squares are of equal lengths, they are conjugates (ref. Lemma 5). Similarly, the shorter squares in these FS-double squares are also conjugates. From Lemma 4, conjugates of sq_1^2 at consecutive locations are possible if sq_1^2 is extended with one of its proper prefixes.

$$\begin{aligned} SQ_1^2 =& (xy)^{p+q}(x)(xy)^p \cdot (xy)^{p+q}(x)(xy)^p \\ =& \underbrace{(\underline{xy})^{p+q}(x)(xy)^{p+q}x}_{sq_1^2} \cdot (\underline{yx})^p(xy)^p \end{aligned}$$

The value of i depends on the longest common prefix of the underlined words in the above structure. Since $xy \neq yx$, the number of conjugates, in this case, is $|lcp(xy, yx)|$. Given any nonempty word u, we can get at most $|u| - 1$ conjugates of a square uu by appending the square by its prefix such that these conjugates start at consecutive locations. So, there can be $|sq_1| - 1$ conjugates that are possible for sq_1^2 and $|lcp(xy, yx)| < |sq_1| - 1$. It shows that the total number of FS-double squares is $|lcp(xy, yx)| + 1$. We know that $q \geq 0$ (see Lemma 2) and this value of i holds for $q > 0$.
It is also necessary to take care of earlier FS-double squares while extending the larger square SQ_1^2. For $q = 0$, the square sq_1^2 repeats if SQ_1^2 is extended with one of its prefixes of size $|x|$ or more. For example, sq_1^2 is a suffix of the word $SQ_1^2.(x)$ as shown below.

$$\begin{aligned} SQ_1^2 =& (xy)^p(x)(xy)^p \cdot (xy)^p(x)(xy)^p \\ SQ_1^2.(x) =& (xy)^p(x)(xy)^p \cdot \boxed{(xy)^p(x)(xy)^p(x)} \implies sq_1^2 \end{aligned}$$

In a word $SQ_1^2.x$, the first location starts with only one rightmost square contradicting the assumption that SQ_1^2 is an FS-double square. So, $i \leq min(|lcp(xy, yx)| + 1, |x|)$. □

3.2. Unequal 2FS Squares

We now consider the possibility that the second FS-double square in a 2FS square is either a γ- or a δ-mate.

Lemma 10. *Given a 2FS square where $|SQ_1| \neq |SQ_2|$. Then, SQ_1^2 and SQ_2^2 cannot be $\gamma-$mates. Moreover, these squares are δ-mates.*

Proof. Let w be a word that begins with two consecutive FS-double squares SQ_1^2 and SQ_2^2 such that SQ_2^2 is a suffix of w. Assume sq_1^2 is the shorter rightmost square beginning at location 1. Similarly, location 2 starts with a shorter square sq_2^2. Referring to the definitions that are mentioned in Subsection 2.1, SQ_2^2 is either a γ or a δ-mate of SQ_1^2. According to Lemma 8, the size of a square starting at location 2 is

equal to either $|sq_1^2|, |SQ_1^2|$ or $2|SQ_1^2|$. If SQ_2^2 is a γ-mate of SQ_1^2, then $|sq_2| = |SQ_1|$. The only possible length of SQ_2 is greater than or equal to $2|SQ_1|$. However, as per the definition of an FS-double square described in Lemma 2, we have $|SQ_2| < 2*|sq_1|$. Here, $|SQ_2| = 2|SQ_1| = 2|sq_2|$ contradicts the condition $|SQ_2| < 2*|sq_1|$. Thus, SQ_1^2 and SQ_2^2 cannot be γ-mates.
Consider the word $w = a((abaaabaabaaabb)(ab)(abaaabaabaaabb))^2$ as a witness to see that δ-mates can start at adjacent locations. □

Theorem 2 (2FS Square). *A 2FS square belongs to one of the following types:*

(a) Equal 2FS square with $|sq_1| = |sq_2|$ and $|SQ_1| = |SQ_2|$, or

(b) Unequal 2FS square with $|sq_1| < |SQ_1|$ and $2|SQ_1| \leq |sq_2| < |SQ_2|$.

Proof. The proof is a direct consequence of Theorem 1, Lemmas 8 and 10. □

Corollary 2. *The two FS-double squares that start at adjacent locations are either α or δ-mates.*

While there is a unique equal 2FS square for a given FS-double square, it turns out that the words with δ-mates have different structures. In other words, the FS-double SQ_2^2 is not unique for SQ_1^2 when SQ_1^2 and SQ_2^2 are δ-mates. This is further explained in detail in Section 4. The results of Lemma 8 and 10 lead to the subsequent lemma.

Lemma 11. *The following statements hold for an unequal 2FS square.*

(a) $|sq_2| \geq 2|SQ_1|$, and

(b) $|SQ_2| > 2|SQ_1|$.

In the next section, we discuss the structure of words having consecutive FS-double squares.

4. Longest Sequence of 2's

We have shown that for a word that starts with consecutive FS-double squares, any two consecutive squares have the structure with either equal or unequal 2FS square. The s_i sequence of such a word has a chain of $2'$s in the beginning. A word w has a sequence of $2'$s if $s_i(w) = s_{i+1}(w) = \cdots = s_j(w) = 2$ where the integers i, j satisfy $1 \leq i < j \leq |w|$. It is possible to extend an FS-double square to get an arbitrarily long sequence of $2'$s. One way to achieve this is described in Lemma 9, where an FS-double square is appended by its prefix. In this case, all the consecutive FS-double squares at the beginning of a word are conjugates. The number of such FS-double squares is finite, and the length of a sequence of $2'$s is limited. However, it is always possible to introduce an unequal 2FS square to increase the length of the sequence of $2'$s. Thus,

we can extend an FS-double square to get a sequence of $2'$s of any desired length by introducing a new equal or an unequal 2FS square.
Under specific conditions, a single letter is added to the FS-double square to introduce a new equal 2FS square. In contrast, the FS-double square is appended by many letters to get a new unequal 2FS square. Let us see some equal and unequal 2FS squares. Next is an example of an equal 2FS square and its s_i sequence. Here, $SQ_1^2 = ((aba)^1(ab)(aba)^1)^2$ and $SQ_2^2 = ((baa)^1(ba)(baa)^1)^2$.

$$\begin{array}{rl} w & = a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a \\ s_i(w) & = 2\ 2\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 0 \end{array}$$

The word, w, if continued to be extended further with the prefix of SQ_1^2, then sq_1^2 repeats after the first location. It reduces the value of s_1 to one. In such words, it is necessary to introduce an unequal 2FS square to further extend the sequence of $2'$s. Unlike equal 2FS squares, the length of an unequal 2FS square varies. There are different ways to extend an FS-double square to get an unequal 2FS square. To elaborate this, we extend an FS-double square in two different ways to get two different unequal 2FS squares. Let $SQ_1^2 = aabaaabaabaaab$ be the FS-double square which is extended to get two unequal 2FS squares w_1 and w_2, where

$$w_1 = a((abaaabaabaaabb)(ab)(abaaabaabaaabb))^2 \tag{9}$$

$$w_2 = a((abaaabaabaaabb)(abaaabaabaaabbab)(abaaabaabaaabb))^2 \tag{10}$$

and the respective s_i sequences are,

$$\begin{array}{rl} w_1 & = \mathrm{a\ a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b\ a\ b\ a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b} \\ & \mathrm{a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b\ a\ b\ a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b} \\ s_i(w_1) & = 2\ 2\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\ & 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0 \end{array}$$

$$\begin{array}{rl} w_2 & = \mathrm{a\ a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b\ a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b\ a\ b\ a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b} \\ & \mathrm{a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b\ a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b\ a\ b\ a\ b\ a\ a\ a\ b\ a\ a\ b\ a\ a\ a\ b\ b} \\ s_i(w_2) & = 2\ 2\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \\ & 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0 \end{array}$$

A word can be extended to get an unequal 2FS square at any location. Moreover, it is possible to yield an unequal 2FS square at a particular location l such that it does not affect the s_i value of another location m where $0 < m < l$. This new 2FS square almost doubles the overall word length, though. So, we investigate the relationship between the length of the longest sequence of $2'$s and the word length. It is evident that the ratio of the longest sequence of $2'$s in a word to its length is higher for equal 2FS squares. The following lemma gives the ratio for the sequence of $2'$s such that any two consecutive FS-double squares in the sequence follow the structure of an equal 2FS square.

Lemma 12 (Longest sequence of $2'$s with Equal 2FS Squares). *Let T be the longest sequence of consecutive FS-double squares in w such that any two consecutive FS-double squares in T are conjugates. Then, $\frac{|T|}{|w|} \leq \frac{1}{7}$.*

Proof. Assume that the first FS-double square SQ_1^2 in T is $((xy)^{p+q}(x)(xy)^p)^2$ where $x, y \in \Sigma^+$ and integers p, q satisfy the relation $p + q \geq 1, q \geq 0$. From Lemma 9, we know that the length of T depends on the values of p and q. The highest value of the ratio $|T|/|w|$ when $q = 0$ is computed below.

$$\frac{|T|}{|w|} = \frac{\min(|lcp(xy, yx)| + 1, (|x| - 1))}{2((p + p + 1)|x| + (p + p)|y|) + |lcp(xy, yx)|}$$
$$= \frac{|x| - 1}{(4p + 3)|x| + 4p|y| - 1} \leq \frac{1}{7}$$

The following equation shows that the ratio $|T|/|w|$ converges to $\frac{1}{7}$ when $q > 0$.

$$\frac{|T|}{|w|} = \frac{|lcp(xy, yx)| + 1}{2((2 + 1 + 1)|x| + (2 + 1)|y|) + |lcp(xy, y, x)|}$$
$$= \frac{|x| + |y| - 2}{8|x| + 6|y|) + |x| + |y| - 2} = \frac{|x| + |y| - 2}{9|x| + 7|y|) - 2} \leq \frac{1}{7}$$

□

The value of $|T|$ is also explored in Lemma 4 of the paper [14] where it is shown that $|T| < \frac{|w|}{2}$. However, the value $|lcp(xy, yx)|$ helps us to show that $\frac{|T|}{|w|}$ is at most $\frac{1}{7}$. The sequence of consecutive FS-double squares obtained in Lemma 12 can be further extended by adding a new unequal 2FS square. So, another way to generate a long sequence of 2's is to start with an FS-double square and extend it to add all possible conjugates of the square. At this point, we can append the word to generate an unequal 2FS square so that the sequence of 2's continues to grow. Thus, a sequence increases either with an equal or an unequal 2FS square. The length of such a sequence in a word with respect to the word length is computed in the following lemma.

Lemma 13 (Longest Sequence of $2'$s with Equal and Unequal 2FS squares). *Let T be the longest sequence of FS-double squares in a word w that contains at least one equal and at least one unequal 2FS square. Then, $\frac{|T|}{|w|} \leq \frac{6}{55}$.*

Proof. The length of T can be increased by adding a new 2FS square. Every new equal 2FS square increments the value of both $|T|$ and $|w|$ by 1. This improves the value of $\frac{|T|}{|w|}$. However, it is not always possible to introduce an equal 2FS square (see Lemma 9) and, therefore, an unequal 2FS square is required to get a longer T. Unlike an equal 2FS square, a new unequal 2FS square decreases the value of $\frac{|T|}{|w|}$. To understand this, suppose an unequal 2FS square begins at location one where SQ_1^2 and SQ_2^2 are two consecutive FS-double squares with shorter squares sq_1^2 and sq_2^2, respectively. Lemma 11 gives the relation $|SQ_2| > 2|SQ_1|$. Accordingly, SQ_1^2 is appended by a word containing at least $2|SQ_1|$ letters to make $s_2 = 2$. Thus, the value of $\frac{|T|}{|w|}$ decreases significantly after introducing an unequal 2FS square. So, we

can obtain the best ratio from the word that has a maximum number of equal 2FS squares and some unequal 2FS squares.
Given a word with $s_1 = s_2 = \cdots = s_i = 2$ such that the location i starts with an FS-double square SQ_i^2. From Lemma 12, the sequence of $2'$s can be extended to get at most $\frac{|SQ_i|}{7}$ new equal 2FS squares. An unequal 2FS square must be introduced to continue the sequence of $2'$s further. The ratio $\frac{|T|}{|w|}$ for the smallest FS-double square $w = (abaab)^2$ is $\frac{1}{10}$. We use the above method to extend w. The respective $\frac{|T|}{|w|}$ obtained after introducing a new unequal 2FS square results into the following sequence.

$$\frac{1 + \frac{10}{7} + 1}{10 + \frac{10}{7} + (10 + 10)}, \frac{1 + \frac{10}{7} + 1 + \frac{20}{7} + 1}{10 + 20 + \frac{10}{7} + \frac{20}{7} + (20 + 20)},$$

$$\frac{1 + \frac{10}{7} + 1 + \frac{20}{7} + 1 + \frac{40}{7} + 1}{10 + 20 + 40 + \frac{10}{7} + \frac{20}{7} + \frac{40}{7} + (40 + 40)}, \ldots$$

The n^{th} term of the above sequence is

$$\frac{(n+1) + \frac{10}{7}\sum_{i=0}^{n-1} 2^i}{\frac{10}{7}\sum_{i=0}^{n-1} 2^i + 10\sum_{i=0}^{n-1} 2^i + 10 * 2^n}$$

Given sequence is a decreasing sequence where the first term is $\frac{6}{55}$ and it converges to $\frac{1}{15}$. □

Theorem 3. *Let T be the longest sequence of $s_i = 2's$ in a word w. Then, $|T| \leq \frac{|w|}{7}$.*

Proof. The computation in Lemmas 12 and 13 shows that the best value of $\frac{|T|}{|w|}$ where T contains either equal length FS-double squares or a combination of equal and unequal 2FS squares. The best value is obtained in the former case, that is, $\frac{1}{7}$. We compare this value with the sequence of T where every two consecutive FS-double squares follow the structure of an unequal 2FS square.
Suppose SQ_1^2 and SQ_2^2 result in an unequal 2FS square at the beginning of a word. Then, $|SQ_2| > 2|SQ_1|$ (see Lemma 11). Thus, to introduce a new unequal 2FS square at location i, it is required to append at least $2|SQ_i|$ letters to the FS-double square SQ_i^2. Allowing only unequal 2FS squares in T, we compute the ratio of the length of the longest sequence of $2'$s in a word to its word length as follows.

$$\frac{1}{10}, \frac{1+1}{10+20}, \frac{1+1+1}{10+20+40}, \ldots, \frac{n}{10 * (2^n - 1)}, \ldots$$

The value of the ratio decreases as n increases and the ratio has the maximum value of $\frac{1}{15}$ for $n = 2$. We ignore the value with $n = 1$ as the sequence will have only one FS-double square. Therefore, $|T| \leq \frac{|w|}{7}$. □

5. Conclusion

We have investigated the ways to get a sequence of FS-double squares by extending a given FS-double square. In this regard, we introduced the term 2FS square for a word starting with two consecutive FS-double squares. A 2FS square is characterized by two types, viz. equal and unequal 2FS square. The former has a single letter added to an existing FS-double square to obtain a new FS-double square. In contrast, the FS-double square is appended by a word of its length (or longer) to yield an unequal 2FS square. Though getting a new equal length FS-double square is easy, we proved that an FS-double square w could be extended to get at most $\frac{|w|}{7}$ equal 2FS squares. In fact, the ratio is also obtained in [3] for the shortest possible words that start with equal length consecutive FS-double squares. On the contrary, we showed that a square w could be extended to have more than $\frac{|w|}{7}$ new FS-double square at consecutive locations by introducing unequal 2FS squares. The overall length of the resulting word increases significantly with the inclusion of such squares. So, we have compared the maximum number of successive FS-double squares in a word with its length. We have found that the ratio of the number of consecutive FS-double squares in an n length word is less than $\frac{6n}{55}$ in the presence of unequal 2FS squares. The best ratio, $\frac{n}{7}$, is possible only with equal 2FS squares.

Recent archived work on the square conjecture [4] came up with a proof validating the square conjecture. If the conjecture is true, the immediate task would be identifying words that pack many distinct squares. A set of similar words is explored in [11] and [17]. These words are considered to represent the lower bound of the square conjecture. Also, they have a variety of structures with a shared property; each such word has a sequence of equal-length squares starting at adjacent locations. So, we believe that the properties of consecutive FS-double squares described in this work will help to construct the best possible words containing the maximum number of distinct squares.

Conflict of interest. The authors declare that they have no conflict of interest.

Data Availability. Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] H. Bai, F. Franek, and W.F. Smyth, *The new periodicity lemma revisited*, Discrete Appl. Math. **212** (2016), 30–36.

https://doi.org/10.1016/j.dam.2016.05.003.

[2] J. Berstel and D. Perrin, *The origins of combinatorics on words*, European J. Combin. **28** (2007), no. 3, 996–1022.
https://doi.org/10.1016/j.ejc.2005.07.019.

[3] F. Blanchet-Sadri and S. Osborne, *Constructing words with high distinct square densities*, arXiv preprint arXiv:1708.06462 (2017), 1–15.

[4] S. Brlek and S. Li, *On the number of squares in a finite word*, arXiv preprint arXiv:2204.10204 (2022), 1–13.

[5] M. Crochemore, *An optimal algorithm for computing the repetitions in a word*, Inform. Process. Lett. **12** (1981), no. 5, 244–250.
https://doi.org/10.1016/0020-0190(81)90024-7.

[6] M. Crochemore and W. Rytter, *Jewels of Stringology: Text Algorithms*, World Scientific, 2002.

[7] A. Deza, F. Franek, and M. Jiang, *A d-step approach for distinct squares in strings*, Annual Symposium on Combinatorial Pattern Matching, Springer, 2011, pp. 77–89.
https://doi.org/10.1007/978-3-642-21458-5_9.

[8] ______, *A computational framework for determining square-maximal Strings.*, Stringology, Citeseer, 2012, pp. 111–119.

[9] A. Deza, F. Franek, and A. Thierry, *How many double squares can a String contain?*, Discrete Appl. Math. **180** (2015), 52–69.
https://doi.org/10.1016/j.dam.2014.08.016.

[10] N.J. Fine and H.S. Wilf, *Uniqueness theorems for periodic functions*, Proc. Amer. Math. Soc. **16** (1965), no. 1, 109–114.

[11] A.S. Fraenkel and J. Simpson, *How many squares can a String contain?*, J. Combin. Theory, Ser. A **82** (1998), no. 1, 112–120.
https://doi.org/10.1006/jcta.1997.2843.

[12] F. Franek and M. Liut, *Computational substantiation of the d-step conjecture for distinct squares revisited*, Prague Stringology Conference 2021, 2021, pp. 41–51.

[13] L. Ilie, *A simple proof that a word of length n has at most 2n distinct squares*, J. Combin. Theory, Ser. A **112**, no. 1, 163–164.
https://doi.org/10.1016/j.jcta.2005.01.006.

[14] ______, *A note on the number of squares in a word*, Theoret. Comput. Sci. **380** (2007), no. 3, 373–376.
https://doi.org/10.1016/j.tcs.2007.03.025.

[15] M. Lothaire, *Applied Combinatorics on Words*, vol. 105, Cambridge University Press, Cambridge, 2005.

[16] F. Manea and S. Seki, *Square-density increasing mappings*, International Conference on Combinatorics on Words, Springer, 2015, pp. 160–169.

[17] M. Patawar and K. Kapoor, *Characterization of dense patterns having distinct squares*, Conference on Algorithms and Discrete Applied Mathematics, Springer, 2021, pp. 397–409.

[18] A. Thierry, *A proof that a word of length n has less than 1.5 n distinct squares*, arXiv preprint arXiv:2001.02996 (2020), 1–30.