New results on pseudosquare avoidance

Tim Ng[1], Pascal Ochem[2], Narad Rampersad[3], and Jeffrey Shallit[1]

[1] School of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
{tim.ng,shallit}@uwaterloo.ca

[2] LIRMM, CNRS, Université de Montpellier, France
ochem@lirmm.fr

[3] Department of Math/Stats, University of Winnipeg, 515 Portage Ave., Winnipeg, MB R3B 2E9, Canada
narad.rampersad@gmail.com

Abstract. We start by considering binary words containing the minimum possible numbers of squares and antisquares (where an antisquare is a word of the form $x\overline{x}$), and we completely classify which possibilities can occur. We consider avoiding $xp(x)$, where p is any permutation of the underlying alphabet, and $xt(x)$, where t is any transformation of the underlying alphabet. Finally, we prove the existence of an infinite binary word simultaneously avoiding all occurrences of $xh(x)$ for *every* nonerasing morphism h and all sufficiently large words x.

1 Introduction

Let x, v be words. We say that v is a *factor* of x if there exist words u, w such that $x = uvw$. For example, `or` is a factor of `word`.

By a *square* we mean a nonempty word of the form xx, like the French word `couscous`. The *order* of a square xx is $|x|$, the length of x. It is easy to see that every binary word of length at least 4 contains a square factor. However, in a classic paper from combinatorics on words, Entringer, Jackson, and Schatz [7] constructed an infinite binary word containing, as factors, only 5 distinct squares: 0^2, 1^2, $(01)^2$, $(10)^2$, and $(11)^2$. This bound of 5 squares was improved to 3 by Fraenkel and Simpson [9]; it is optimal. For some other constructions also achieving the bound 3, see [15,14,10,2].

Instead of considering squares, one could consider *antisquares*: these are binary words of the form $x\overline{x}$, where $\overline{x}$ is a coding that maps $0 \to 1$ and $1 \to 0$. For example, `01101001` is an antisquare. (They should not be confused with the different notion of antipower recently introduced by Fici, Restivo, Silva, and Zamboni [8].) Clearly it is possible to construct an infinite binary word that avoids all antisquares, but only in a trivial way: the only such words are $0^\omega = 000\cdots$ and $1^\omega = 111\cdots$. Similarly, the only infinite binary words with exactly one antisquare are 01^ω and 10^ω. However, it is easy to see that every word in $\{1000, 10000\}^\omega$ has exactly two antisquares — namely 01 and 10 — and hence there are infinitely many such words that are aperiodic.

Several writers have considered variations on these results. For example, Blanchet-Sadri, Choi, and Mercaş [3] considered avoiding large squares in partial words. Chiniforooshan, Kari, and Zhu [4] studied avoiding words of the form $x\theta(x)$, where θ is an antimorphic involution. Their results implicitly suggest the general problem of simultaneously avoiding what we might call *pseudosquares*: patterns of the form xx', where x' belongs to some (possibly infinite) class of modifications of x.

This paper has two goals. First, for all integers $a, b \geq 0$ we determine whether there is an infinite binary word having at most a squares and b antisquares. If this is not possible, we determine the length of the longest finite binary word with this property.

Second, we apply our results to discuss the simultaneous avoidance of xx', where x' belongs to some class of modifications of x. We consider three cases:

(a) where $x' = p(x)$ for a permutation p of the underlying alphabet;
(b) where $x' = t(x)$ for a transformation t of the underlying alphabet; and
(c) where $x' = h(x)$ for an arbitrary nonerasing morphism.

In particular, we prove the existence of an infinite binary word that avoids $xh(x)$ simultaneously for all nonerasing morphisms h and all sufficiently long words x.

2 Simultaneous avoidance of squares and antisquares

We are interested in binary words where the number of distinct factors that are squares and antisquares is bounded. More specifically, we completely solve this problem determining in every case the length of the longest word having at most a distinct squares and at most b distinct antisquares. Our results are summarized in the following table. If (one-sided) infinite words are possible, this is denoted by writing ∞ for the length.

The results in the first two columns and first three rows (that is, for $a \leq 2$ and $b \leq 1$) are very easy. We first explain the first two columns:

Proposition 1.

(a) For $a \geq 0$, the longest binary word with a squares and 0 antisquares has length $2a + 1$.
(b) For $a \geq 0$, the longest binary word with a squares and 1 antisquare has length $2a + 2$.

Proof.

(a) If a binary word has no antisquares, then in particular it has no occurrences of either 01 or 10. Thus it must contain only one type of letter. If it has length $2a + 2$, then it has $a + 1$ squares, of order $1, 2, \ldots, a + 1$. If it has length $2a + 1$, it has a squares. So $2a + 1$ is optimal.

$a \backslash b$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	···
0	1	2	3	3	3	3	3	3	3	3	3	3	3	3	···
1	3	4	7	7	7	7	7	7	7	7	7	7	7	7	···
2	5	6	11	11	11	11	12	12	12	13	15	18	18	18	···
3	7	8	15	15	15	20	20	20	24	29	34	53	98	∞	···
4	9	10	19	19	27	31	45	56	233	∞	∞	∞	∞	···	
5	11	12	27	27	40	∞	∞	∞	∞	···					
6	13	14	35	38	313	∞	···								
7	15	16	45	∞	∞	···									
8	17	18	147	∞	···										
9	19	20	∞	···											
10	21	22	∞	···											
⋮															

Fig. 1. Length of longest binary word having at most a squares and b antisquares

(b) If a length-n binary word w has only one antisquare, this antisquare must be either 01 or 10; without loss of generality, assume it is 01. Then w is either of the form $0^{n-1}1$ or 01^{n-1}. Such a word clearly has $\lfloor (n-1)/2 \rfloor$ squares.

We next explain the first three rows: if a binary word has no squares, its length is clearly bounded by 3, as we remarked earlier. If it has one square, a simple argument shows it has length at most 7. Finally, if it has two squares, already Entringer, Jackson, and Schatz [7, Thm. 2] observed that it has length at most 18.

For all the remaining finite entries, we obtained the result through the usual backtrack search method, and we omit the details.

In what follows, we provide the lexicographically least binary words achieving the "important" bounds in Figure 1.

$(3, 12)$:
0010001100101110001011001110001100101110001011000
1110010111000101100111000110010111000101100111011

$(4, 8)$:
000010000001100001011000000110001011000000101110001011000001011
100000101100000011000101100000010111000101100000011000101100000
101110000001011000000110001011000000101110001011000010111000000
1011000000110001011000010111000101100000011000101100001000

$(5, 4)$: 0001000000101000000001010100000001010000100

(6, 4) : 0000100000010100000001101000000010100001101000000010100000011010
0000101000011010000000101000001101000000101000011010000000010100
0000110100000101000011010000000101000001101000000010100000011010
0000101000011010000000101000001101000000101000011010000000010100
000011010000010100001101000000010100000110100000000110101111010

(7, 2): 000001000000010100000010000101000000010000101

(8, 2) : 0010000101000100010001010000010001000100000101000
1000100010100000100010000010100010001010000010001
00010000001010001000100010100000100010001000001010

It now remains to prove the results labeled ∞. First, we introduce some morphisms. Let the morphisms $h_{3,13}$, $h_{4,9}$, $h_{5,5}$, $h_{7,3}$, $h_{9,2}$ be defined as follows:

(a) $h_{3,13}$: 0 → 0010110011100011
1 → 001011000111
2 → 00101110

(b) $h_{4,9}$: 0 → 0000101110000011000010110000011000101100001011100010110
1 → 0000101110000011000010110000011000101100000101110001011
2 → 0000101110000011000010110000010111000101100000110001011
This is a 55-uniform morphism.

(c) $h_{5,5}$: 0 → 1010000001011000010100001101011000001
1 → 1010000001011000001101011000010100001
2 → 101000000101000011000000101000000110000
This is a 36-uniform morphism.

(d) $h_{7,3}$: 0 → 0100100100001010000
1 → 01001001000001
2 → 0100100101000

(e) $h_{9,2}$: 0 → 0001000100000001000101
1 → 0000010001000100000101
2 → 0000001000100000010100
This is a 22-uniform morphism.

Theorem 1. *Let* $\mathbf{w}$ *be an infinite squarefree sequence over the alphabet* $\{0, 1, 2\}$. *Then* $h_{a,b}(\mathbf{w})$ *contains exactly* a *squares and* b *antisquares. More precisely*

(a) $h_{3,13}(\mathbf{w})$ *contains the squares* 0^2*,* 1^2*, and* $(01)^2$ *and the antisquares* 01, 10, 0011, 0110, 1001, 1100, 000111, 001110, 011100, 100011, 110001, 111000, *and* 10010110.
(b) $h_{4,9}(\mathbf{w})$ *contains the squares* 0^2, 1^2, $(00)^2$, *and* $(01)^2$ *and the antisquares* 01, 10, 0011, 0110, 1100, 011100, 110001, 111000, *and* 1110000011.
(c) $h_{5,5}(\mathbf{w})$ *contains the squares* 0^2, 1^2, $(00)^2$, $(01)^2$, *and* $(10)^2$ *and the antisquares* 01, 10, 0011, 0110, *and* 1100.
(d) $h_{7,3}(\mathbf{w})$ *contains the squares* 0^2, $(00)^2$, $(01)^2$, $(10)^2$, $(001)^2$, $(010)^2$, *and* $(100)^2$ *and the antisquares* 01, 10, *and* 1001.
(e) $h_{9,2}(\mathbf{w})$ *contains the squares* 0^2, $(00)^2$, $(01)^2$, $(10)^2$, $(000)^2$, $(0001)^2$, $(0010)^2$, $(0100)^2$, *and* $(1000)^2$ *and the antisquares* 01 *and* 10.

Proof. Let h be any of the morphisms above. We first show that large squares are avoided. The h-images of the letters have been ordered such that $|h(0)| \geq |h(1)| \geq |h(2)|$. A computer check shows that for every letter i and every ternary word $\mathbf{w}$, the factor $h(i)$ appears in $h(\mathbf{w})$ only as the h-image of i. Another computer check shows that for every ternary squarefree word $\mathbf{w}$, the only squares uu with $|u| \leq 2|h(0)| - 2$ that appear in $h(w)$ are the ones we claim. If $h(w)$ contains a square uu with $|u| \geq 2|h(0)| - 1$, then u contains the full h-image of some letter. Thus, uu is a factor of $h(avbvc)$ with a, b, c single letters and v a nonempty word. Moreover, $a \neq b$ and $b \neq c$, since otherwise $avbvc$ would contain a square. It follows that $u = ph(v)s$, so that p is a suffix of $h(a)$, $h(b) = sp$, and s is a prefix of $h(c)$. Thus, $h(abc)$ contains the square $psps$ with period $|ps| = |h(b)|$. Since $5 < |h(2)| \leq |h(b)| \leq |h(0)| < 2|h(0)| - 2$, this contradicts our computer check, which rules out squares with period at least 5 and at most $2|h(0)| - 2$.

To show that large antisquares are avoided, it suffices to exhibit a factor f such that f is uniformly recurrent in $h(\mathbf{w})$ and $\overline{f}$ is not a factor of $h(\mathbf{w})$. We use $f = 0101$ for $h_{3,13}$ and $f = 0^4$ for the other morphisms.

Remark 1. The uniform morphisms were found as follows: for increasing values of q, our program looks (by backtracking) for a binary word of length $3q$ corresponding to the image $h(012)$ of 012 by a suitable q-uniform morphism h. Given a candidate h, we check that $h(w)$ has at most a squares and b antisquares for every squarefree word $\mathbf{w}$ up to some length. Standard optimizations are applied to the backtracking. Squares and antisquares are counted naively (recomputed from scratch at every step), which is sufficient since the morphisms found are not too large.

Remark 2. The morphisms $h_{3,13}$ and $h_{7,3}$ are not uniform. However, we can construct uniform morphisms with the same properties as follows. Let m be the 18-uniform squarefree morphism given by

$$\begin{aligned} 0 &\to 021012102012021201 \\ 1 &\to 021012102120210201 \\ 2 &\to 021012102120102012 \,. \end{aligned}$$

Notice that $m(0)$, $m(1)$, and (2) contain 6 occurrences of each letter. So the 216-uniform morphism $h'_{3,13} = h_{3,13} \circ m$ is such that $h'_{3,13}(\mathbf{w})$ and $h_{3,13}(\mathbf{w})$ contain the same squares and antisquares. Similarly, for binary words with the same squares and antisquares as $h_{7,3}(w)$, we can use the 276-uniform morphism $h_{7,3} \circ m$. However in this case, we have found the following smaller morphism, which is 29-uniform.

$$
\begin{aligned}
0 &\rightarrow 00101000010010010100000101001 \\
1 &\rightarrow 00101000010010010000101001000 \\
2 &\rightarrow 00101000010010010000101000001 \ .
\end{aligned}
$$

Corollary 1. *There exists an infinite binary word having at most ten distinct squares and antisquares as factors, but the longest binary word having nine or fewer distinct squares and antisquares is of length* 45*.*

Remark 3. A word of length 45 with a total of nine distinct squares and antisquares is

$$000001000000010100000001000010100000000010000101.$$

Corollary 2. *Every infinite word having at most ten distinct squares and antisquares has critical exponent at least* 5*, and there is such a word having* 5*-powers but no powers of higher exponent.*

Proof. By the usual backtracking approach, we can easily verify that the longest finite word having at most ten distinct antisquares, and critical exponent < 5 is of length 57. One such example is

$$010001010000100100100001010010010100001001001000010100010.$$

On the other hand, if $\mathbf{w}$ is any squarefree ternary infinite word, then from above we know that the only possible squares that can occur in $h_{5,5}(\mathbf{w})$ are of the form x^2 for $x \in \{0, 1, 00, 01, 10\}$. It is now easy to verify that the largest power of 0 that occurs in $h_{5,5}(\mathbf{w})$ is 0^5; the largest power of 1 that occurs is 1^2; the largest power of 01 that occurs is $(01)^{5/2}$; and the largest power of 10 that occurs is $(10)^{5/2}$.

Proposition 2. *Every infinite cubefree binary word has a total of at least 23 distinct squares and antisquares.*

Proof. By the usual backtracking method.

Remark 4. In the final version of this paper, we plan to provide the optimal bound.

3 Pseudosquare avoidance

In this section we discuss avoiding xx' where x' belongs to some large class of modifications of x'. This is in the spirit of previous results [16,5,12], where one is interested in avoiding factors of low Kolmogorov complexity. The problems we study are not quite so general, but our results are effective, and we obtain explicit bounds.

3.1 Avoiding pseudosquares for permutations

Here we are interested in avoiding patterns of the form $xp(x)$, for *all* codings p that are permutations of the underlying alphabet. Of course, this is impossible for words of length ≥ 2 strictly as stated, since every word of length 2 is of the form $ap(a)$ where p is the permutation sending the letter a to $p(a)$. Thus it is reasonable to ask about avoiding $xp(x)$ for all words x of length $\geq n$. Our first result shows this is impossible for $n = 2$.

Theorem 2. *For all finite alphabets Σ, and for all words w of length ≥ 10 over Σ, there exists a permutation p of Σ and a factor of w of the form xx', where $x' = p(x)$, and $|x| \geq 2$.*

Proof. Using the usual tree-traversal technique, where we extend the alphabet size at each length extension.

We now turn to the case of larger n. For $n \geq 3$, we can avoid all factors of the form $xp(x)$ over the binary alphabet. Of course, this case is particularly simple, since there are only two permutations of the alphabet: the identity permutation that leaves letters invariant, and the map $x \to \overline{x}$, which changes 0 to 1 and vice versa.

Theorem 3. *There exists an infinite word* $\mathbf{w}$ *over the binary alphabet $\Sigma_2 = \{0, 1\}$ that avoids xx and $x\overline{x}$ for all x with $|x| \geq 3$.*

Proof. We can use the morphism $h_{5,5}$ in Theorem 1 (c). Alternatively, a simpler proof comes from the fixed point of the morphism

$$\begin{aligned} 0 &\to 01 & 1 &\to 23 \\ 2 &\to 24 & 3 &\to 51 \\ 4 &\to 06 & 5 &\to 01 \\ 6 &\to 74 & 7 &\to 24 \end{aligned}$$

followed by the coding $n \to n \bmod 2$. We can now use `Walnut` [13] to verify that the resulting 2-automatic word has the desired property. This word has exactly 5 distinct squares:

$$0^2, 1^2, (00)^2, (01)^2, (10)^2,$$

and exactly 6 distinct antisquares:

$$01, 10, 0011, 0110, 1001, 1100.$$

3.2 Avoiding pseudosquares for transformations

In the previous subsection we considered permutations of the alphabet. We now generalize this to *transformations* of the alphabet, or, in other words, to arbitrary codings (letter-to-letter morphisms).

Theorem 4.

(a) *For all finite alphabets Σ, and all words w of length ≥ 31 over Σ, there exists a transformation $t : \Sigma^* \to \Sigma^*$ such that w contains a factor of the form $xt(x)$ for $|x| \geq 3$.*
(b) *For all finite alphabets Σ, and all words w of length ≥ 16 over Σ, there exists a transformation t of Σ such that w contains a factor of the form xx', where $x' = t(x)$ or $x = t(x')$ and $|x| \geq 3$.*

Proof. Using the usual tree-traversal technique, where we extend the alphabet size at each length extension.

We now specialize to the binary alphabet. This case is particularly simple, since in addition to the two permutations of the alphabet, the only other transformations are the ones sending both $0, 1$ to a single letter (either 0 or 1).

Theorem 5. *There exists an infinite word* $\mathbf{w}$ *over the binary alphabet $\Sigma_2 = \{0, 1\}$ avoiding 0^4, 1^4, and xx and $x\overline{x}$ for every x with $|x| \geq 4$. In other words,* $\mathbf{w}$ *avoids both $xt(x)$ and $t(x)x$ for $|x| \geq 4$ and all transformations t.*

Proof. Use the fixed point of the morphism

$$0 \to 01 \qquad 1 \to 23$$
$$2 \to 45 \qquad 3 \to 21$$
$$4 \to 23 \qquad 5 \to 42$$

followed by the coding $n \to \lfloor n/3 \rfloor$. The result can now easily be verified with `Walnut`.

3.3 Avoiding pseudosquares with morphic images

In this subsection we consider simultaneously avoiding all patterns of the form $xh(x)$, for all morphisms h defined over $\Sigma_k = \{0, 1, \ldots, k-1\}$. Clearly this is impossible if h is allowed to be erasing (that is, some images are allowed to be empty), or if x consists of a single letter. So once again we consider the question for sufficiently long x.

For this version of the problem, it is particularly hard to obtain experimental data, because the problem of determining, given x and y, whether there is a morphism h such that $y = h(x)$, is NP-complete [1,6].

Theorem 6. *No infinite word over a finite alphabet avoids all factors of the form $xh(x)$, for all nonerasing morphisms h, with $|x| = 4$.*

Proof. Let $\mathbf{w}$ be a potential counter-example to Theorem 6. Without loss of generality, we can assume that $\mathbf{w}$ is uniformly recurrent (see, e.g., [11, Lemma 2.4]). We use a, b, and c to denote distinct letters and u and v to denote non-empty finite words.

A word is called *basic* if it is of the form au such that $|u| = 3$ and u does not contain the letter a. Suppose, to get a contradiction, that $\mathbf{w}$ contains a basic factor. Since u is recurrent, the factor au extends to $auvu$, which is a forbidden occurrence of $xh(x)$. So $\mathbf{w}$ avoids basic factors.

Suppose, to get a contradiction, that $\mathbf{w}$ contains a 4-power v^4. Then $\mathbf{w}$ contains a factor uv^4 with $|u| = 4$, which is a forbidden occurrence of $xh(x)$. So $\mathbf{w}$ is 4-power free.

Suppose, to get a contradiction, that $\mathbf{w}$ contains a factor aaa. Since $\mathbf{w}$ does not contain the 4-power $aaaa$, then $\mathbf{w}$ must contain $baaa$, which is a basic factor. So $\mathbf{w}$ avoids aaa.

Suppose, to get a contradiction, that $\mathbf{w}$ contains three consecutive distinct letters abc. To avoid a basic factor, abc must extend to $abca$. Then $abca$ must extend to $abcab$, and so on. Thus $\mathbf{w}$ must contain the 4-power $(abc)^4$. So $\mathbf{w}$ avoids abc.

Since $\mathbf{w}$ avoids aaa, abc, and the basic factor $abbc$, it must be that $\mathbf{w}$ is a binary word.

Suppose, to get a contradiction, that $\mathbf{w}$ contains a factor $abaabb$. Since bb is recurrent, the factor $abaab$ extends to $abaabubb$, which is a forbidden occurrence of $xh(x)$. So $\mathbf{w}$ avoids $abaabb$,

Finally, $\mathbf{w}$ avoids $ababbaba$ and $ababaabaab$, which are forbidden occurrences of $xh(x)$. Using the usual tree-traversal technique, we check that no infinite 4-power free binary word avoids aaa, $abaabb$, $ababbaba$, and $ababaabaab$. Thus, $\mathbf{w}$ does not exist.

Theorem 7. *There exists an infinite binary word that avoids all factors of the form $xh(x)$ and $h(x)x$, for all nonerasing binary morphisms h, with $|x| \geq 5$.*

Proof. Let $\mathbf{w} = m(\mathbf{t})$, where $\mathbf{t}$ is any ternary squarefree word $\mathbf{t}$ and m is the 57-uniform morphism given below.

$$0 \to 101000110010100110001011001010110001010100011001011000110$$
$$1 \to 101000110010100110001010110001101001100010101000110101001$$
$$2 \to 101000110010100110001010100011010011000101011000110101001$$

We use u and v to denote non-empty words and a to denote a letter. We will need the following properties of $\mathbf{w}$.

(a) The only squares occurring in $\mathbf{w}$ are 00, 11, 0101, and 1010.
(b) $\mathbf{w}$ does not contain any factor uvu with $|u| \geq 15$ and $|v| \leq 6$.
(c) $\mathbf{w}$ does not contain any of the following factors: 111, 110011, 101011001, 100110101, 0100010, 00100.
(d) Every factor of $\mathbf{w}$ of length 13 contains 000, 001100, 010100110, or 011001010.
(e) Every factor of $\mathbf{w}$ of length 12 contains 0101 or 1010.

(f) Every factor of $\mathbf{w}$ of length at least 5, except 00010 and 01000, contains 0011, 1100, 0101, 1010, 0110, 1001, or 10001.

The proofs of (a) and (b) are similar to the proof of Theorem 1. The other properties can be checked by inspecting factors of $\mathbf{w}$ with bounded length. The following cases show that $\mathbf{w}$ contains no factor of the form $xh(x)$ or $h(x)x$ with $|x| \geq 5$.

- We rule out $h(0) = h(1)$, as $h(x)$ contains $h(0)^5$, which contradicts (a).

- We rule out $h(a) = a$, as $xh(x) = h(x)x = xx$ is a square with period at least 5, which contradicts (a).

- We rule out $|x| \geq 13$: By (e), x contains 0101 or 1010. By (a), $\mathbf{w}$ contains no square with period at least 3, which forces $|h(0)| = |h(1)| = 1$. By the previous cases, the only remaining possibility is $h(a) = \overline{a}$. By (d), x contains a factor $v \in \{000, 001100, 010100110, 011001010\}$. Thus, $h(x)$ contains the factor $\overline{v} \in \{111, 110011, 101011001, 100110101\}$, which contradicts (c).

- We rule out that x contains $a\overline{a}a\overline{a}$ or $aa\overline{a}\overline{a}$: Notice that both letters 0 and 1 are contained in a square. By (a), $|h(0)| \leq 2$ and $|h(1)| \leq 2$. A computer check shows that $\mathbf{w}$ contains no factor of the form $xh(x)$ or $h(x)x$ such that $5 \leq |x| \leq 12$, $|h(0)| \leq 2$, and $|h(1)| \leq 2$.

- We rule out that x contains $a\overline{a}\overline{a}a$ or 10001: in every case $h(x)$ contains a factor uvu such that v is square or a cube. By (a), $\min\{|h(0)|, |h(1)|\} \leq 2$ and $|v| \leq 6$. By (b), this means that $\max\{|h(0)|, |h(1)|\} \leq 14$. Again, we only have to consider the case $5 \leq |x| \leq 12$, so that $xh(x)$ and $h(x)x$ have bounded length and can be ruled out by computer check.

- By (f), the only remaining possibilities are $x = 00010$ and $x = 01000$. Notice that if $h(000)$ is in $\mathbf{w}$, then $h(0) \in \{0, 01, 10\}$. Suppose that $x = 00010$. An occurrence of $h(x)x$ contains 0000 if $h(0) \in \{0, 10\}$ and contains 0100010 if $h(0) = 01$. This contradicts (a) and (c), respectively. An occurrence of $xh(x)$ contains 00100 if $h(0) \in \{0, 01\}$ and contains 10101010 if $h(0) = 10$. This contradicts (c) and (a), respectively. The case $x = 01000$ is symmetrical by reversal.

4 Future work

In the future we could consider similar questions for abelian avoidability problems.

Acknowledgments

We thank the referees for many useful suggestions.

References

1. Angluin, D.: Finding patterns common to a set of strings. J. Comput. System Sci. **21**, 46–62 (1980)
2. Badkobeh, G., Crochemore, M.: Fewest repetitions in infinite binary words. RAIRO Inform. Théor. App. **46**, 17–31 (2012)
3. Blanchet-Sadri, F., Choi, I., Mercaş, R.: Avoiding large squares in partial words. Theoret. Comput. Sci. **412**, 3752–3758 (2011)
4. Chiniforooshan, E., Kari, L., Xu, Z.: Pseudopower avoidance. Fund. Inform. **114**(1), 55–72 (2012)
5. Durand, B., Levin, L., Shen, A.: Complex tilings. J. Symbolic Logic **73**, 593–613 (2008)
6. Ehrenfeucht, A., Rozenberg, G.: Finding a homomorphism between two words is NP-complete. Inform. Process. Lett. **9**, 86–88 (1979)
7. Entringer, R.C., Jackson, D.E., Schatz, J.A.: On nonrepetitive sequences. J. Combin. Theory. Ser. A **16**, 159–164 (1974)
8. Fici, G., Restivo, A., Silva, M., Zamboni, L.Q.: Anti-powers in infinite words. J. Combin. Theory. Ser. A **157**, 109–119 (2018)
9. Fraenkel, A.S., Simpson, J.: How many squares must a binary sequence contain? Electronic J. Combinatorics **2**, #R2 (1995)
10. Harju, T., Nowotka, D.: Binary words with few squares. Bull. European Assoc. Theor. Comput. Sci. (89), 164–166 (2006)
11. Luca, A.d., Varricchio, S.: Finiteness and iteration conditions for semigroups. Theoret. Comput. Sci. **87**, 315–327 (1991)
12. Miller, J.S.: Two notes on subshifts. Proc. Amer. Math. Soc. **140**, 16171622 (2012)
13. Mousavi, H.: Automatic theorem proving in `Walnut` (2016), available at http://arxiv.org/abs/1603.06017
14. Ochem, P.: A generator of morphisms for infinite words. RAIRO Inform. Théor. App. **40**, 427–441 (2006)
15. Rampersad, N., Shallit, J., w. Wang, M.: Avoiding large squares in infinite binary words. Theoret. Comput. Sci. **339**, 19–34 (2005)
16. Rumyantsev, A.Y., Ushakov, M.A.: Forbidden substrings, Kolmogorov complexity and almost periodic sequences. In: STACS 2006, Lecture Notes in Computer Science, vol. 3884, pp. 396–407. Springer-Verlag (2006)