A 2D non-overlapping code over a q-ary alphabet

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

A 2D non-overlapping code over a q-ary alphabet

**Elena Barcucci · Antonio Bernini ·
Stefano Bilotta · Renzo Pinzani**

the date of receipt and acceptance should be inserted later

Abstract We define a set of matrices over a finite alphabet where all possible overlaps between any two matrices are forbidden. The set is also enumerated by providing some recurrences counting particular classes of restricted words. Moreover, we analyze the cardinality of the set according to certain parameters related to the construction of the matrices.

Keywords: Bidimensional codes, Non-overlapping matrices, restricted words.

2010 Mathematics Subjects Classification: 68R15; 94B25; 05A15.

1 Introduction

In the present paper, moving from [6], we define a set of *non-overlapping* matrices over a finite alphabet having cardinality $q \geq 2$. Roughly speaking, two matrices (possibly the same) are non-overlapping (self non-overlapping) when it is not possible to shift one on the other one in such a way that the entries of the intersection match (note that the notion of self non-overlapping matrices coincides with the one of *unbordered pictures* proposed in [3]).

More precisely, given two matrices A and B, we can imagine to make a rigid movement of B on A such that B glides on A. At the end of each slipping, which can be geometrically interpreted as a translation in a given direction on the plane, a (non empty) common area is formed. This common area can be seen as the usual intersection between the two rectangular arrays containing the entries of A and B, which is, in turn, a rectangular array constituted by a finite number of 1×1 cells of the discrete plane. Each cell of the common area contains an entry of A and an entry of B. If in each cell of the common area the entry of A coincides with the entry of B, then A and B are overlapping matrices. On the contrary, if

E. Barcucci · A. Bernini · S. Bilotta · R. Pinzani
Dipartimento di Matematica e Informatica "U. Dini", Universitá degli Studi di Firenze,
Viale G.B. Morgagni 65, 50134 Firenze, Italy.
E-mail: elena.barcucci@unifi.it, antonio.bernini@unifi.it, stefano.bilotta@unifi.it, renzo.pinzani@unifi.it

for any translation we never find an overlapping common area, then A and B are non-overlapping matrices.

The matrices of the set we are going to define are required to have some fixed entries on their frame while the rows are subjected to some constraints, including pattern avoidance requirements.

The notion of pattern is one of the most studied in combinatorics, starting from [24] where it was introduced for permutations. Successively it has been considered also in the context of many other combinatorial structures, such as set partitions [23,29], words or binary words [12,13,17], trees [28] and lattice paths [10]. Some of these structures have also been endowed with a partial order structure obtaining in some cases several interesting results [4,11,15].

For our purpose we are going to consider words avoiding consecutive patterns (factors). In particular, each row of the matrices of the set must avoid some specified consecutive patterns of length k depending on the entries used to fill (part of) the frame of the matrices. In the case of the binary alphabet, these consecutive forbidden patterns are simply 0^k and 1^k (see [6]). A first straightforward extension to a q-ary alphabet Σ ($q \geq 2$) consists in choosing two symbols of Σ both for the entries of the frame of the matrices and, consequently, for the forbidden patterns. Actually, a deeper inspection of the matter reveals the possibility of giving a more general definition of the matrices. This new approach leads to a set of non-overlapping matrices whose cardinality, in some cases, turns out to be greater than the one of the set obtained with the above mentioned simple extension. In particular, here a number $p \geq 2$ of symbols of Σ (so $p \leq q$) can be considered for the entries of the frame of the matrices.

The present paper can be seen as an extension to the bidimensional case of the theory of non-overlapping strings. In literature, in the linear case, sets of non-overlapping strings are said *cross-bifix-free sets* (or *codes*). Recently, the effort of the researches has been devoted to the definition of sets having cardinality as large as possible [7,8,14,16,18], once the length n of the strings and the cardinality q of the alphabet are fixed. In our work we follow the same trend. In particular, once fixed the dimension $m \times n$ of the matrices, we give some results about the cardinality of our sets depending on k, p and q.

The increasing interest for digital image processing validates several recent works on the theory of two-dimensional languages [1,2,26,27], which is an additional area where our work can be placed, even if most of the classical matters on formal languages are not here considered. Moreover, our set can be seen as a bidimensional code which is a set of matrices X over Σ such that any matrix over Σ has only one tiling decomposition involving elements of X (see [2]).

The paper is organized as follows: after the definition of the main set (Section 2), we give the enumeration of it by means of recurrence relations and their generating functions, providing also some results on the asymptotics of the number of matrices (Section 3). In Section 4 we give some considerations about the maximal size of our set depending on specific parameters. We conclude (Section 5) with some open problems and further developments.

2 A set of non-overlapping q-ary matrices

Let $\Sigma = \{0, 1, \ldots, q-1\}$ be a finite alphabet with q different symbols, with $q \geq 2$. Let $\mathcal{M}_{m\times n,q}$ be the set of all the matrices with m rows and n columns over Σ. Two distinct matrices $A, B \in \mathcal{M}_{m\times n,q}$ are said *non-overlapping* if there are no translations in any direction of A on B (of B on A) such that the entries of the intersection match. In the case $A = B$, the matrix is said *self non-overlapping* (or *unbordered*). The above concept can be also formally defined as [6] involving block matrices. Nevertheless, here we do not report the related definitions.

Fixed the dimension $m \times n$ of the matrices, we now define a possible non-overlapping set where the matrices have a particular structure involving some of the entries on the frame of the matrix. More precisely, two symbols of Σ are chosen and used to fill part of the frame of the matrix. Without loss of generality we consider 0 and 1. The following definition moves from the one provided in [6] for the binary alphabet, and here it is adapted for our purpose to q-ary alphabet.

Definition 1 Let $2 \leq k \leq \lfloor \frac{n}{2} \rfloor$. We denote $\mathcal{S}^{(k)}_{m\times n,q} \subset \mathcal{M}_{m\times n,q}$ the set of the matrices $A = (a_{i,j})$ satisfying the following conditions:

- $A_1 = 0^{k-1}1w_101^{k-1}$, where $v_1 = 1w_10$ is a q-ary string of length $n - 2k + 2$ avoiding both 0^k and 1^k;
- for $i = 2, \ldots, m-1$, $A_i = w_i1 = v_i$, where v_i is a q-ary string of length n avoiding both 0^k and 1^k;
- $A_m = 0^k v_m 1^k$, where v_m is a q-ary string of length $n - 2k$ avoiding both 0^k and 1^k.

(With A_1, A_i and A_m we denote the first, the i-th and the m-th row of the array A.)

In other words, some entries on the frame of a matrix in $\mathcal{S}^{(k)}_{m\times n,q}$ are fixed. For example, a matrix in $\mathcal{S}^{(3)}_{6\times 10,4}$ is represented in Figure 1. Note that the patterns 222 and 333 are allowed in v_i, while 000 and 111 are forbidden in v_i, with $i = 1, 2, \ldots, 6$.

$$\begin{pmatrix}
\mathbf{0} & \mathbf{0} & \mathbf{1} & 1 & 3 & 2 & 2 & \mathbf{0} & \mathbf{1} & \mathbf{1} \\
2 & 3 & 1 & 1 & 2 & 2 & 2 & 2 & 0 & \mathbf{1} \\
3 & 3 & 3 & 2 & 1 & 2 & 1 & 2 & 1 & \mathbf{1} \\
1 & 1 & 2 & 0 & 0 & 2 & 3 & 1 & 3 & \mathbf{1} \\
0 & 1 & 0 & 3 & 3 & 2 & 1 & 0 & 2 & \mathbf{1} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & 2 & 1 & 0 & 2 & \mathbf{1} & \mathbf{1} & \mathbf{1}
\end{pmatrix}$$

Fig. 1: A matrix in $\mathcal{S}^{(3)}_{6\times 10,4}$

The following proposition holds.

Proposition 1 *The set* $\mathcal{S}^{(k)}_{m\times n,q} \subset \mathcal{M}_{m\times n,q}$ *is non-overlapping, for each* k *with* $2 \leq k \leq \lfloor \frac{n}{2} \rfloor$ *and* $m \geq 2$.

Proof. We proceed ad absurdum. If two matrices belonging to $\mathcal{S}^{(k)}_{m\times n,q}$ are overlapping, then we have two cases. In the first case, as in the top of Figure 2, one of the two matrices has to contain a forbidden pattern of length k against the hypothesis that it is in $\mathcal{S}^{(k)}_{m\times n,q}$. The second case, pictured in the bottom of Figure 2, occurs when the fixed entries of the frames of the two matrices overlap. In this case at least one of these entries is not defined.

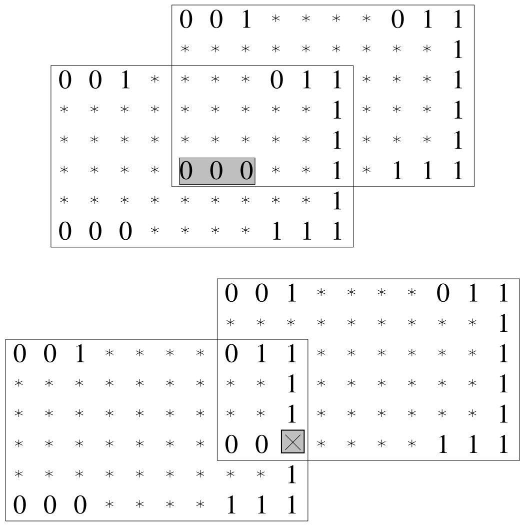

Fig. 2: In the top: grey entries are forbidden; in the bottom: grey entry contains different values and it is not defined

■

We now propose a possible extension of the previous definition leading to a new set of matrices involving more than two values for the fixed entries on the frame of the matrices.

Definition 2 Let $P = \{0, 1, \dots, p-1\} \subseteq \Sigma$, with $2 \le p \le q$. Let $\{P_1, P_2\}$ a partition of P (i.e. $P_1 \cup P_2 = P$ and $P_1 \cap P_2 = \emptyset$). Without loss of generality we can assume $P_1 = \{0, 1, 2, \dots, j-1\}$ and $P_2 = \{j, j+1, \dots, p-1\}$. For each $(\alpha, \beta) \in P_1 \times P_2$ and $2 \le k \le \left\lfloor \frac{n}{2} \right\rfloor$ we denote

$$\mathcal{S}^{(k,p)}_{m\times n,q} = \bigcup_{(\alpha,\beta)} S^{(k,p)}_{m\times n,q}(\alpha, \beta)$$

where $S^{(k,p)}_{m\times n,q}(\alpha, \beta)$ is the set of matrices $A = (a_{i,j})$ satisfying the following conditions:

- $A_1 = \alpha^{k-1}\beta w_1 \alpha \beta^{k-1}$, where $v_1 = \beta w_1 \alpha$ is a q-ary string of length $n - 2k + 2$ avoiding all the patterns $0^k, 1^k, 2^k, \dots, (p-1)^k$;

$$\begin{pmatrix} 0&0&1&*&*&*&0&1&1\\ *&*&*&*&*&*&*&*&1\\ *&*&*&*&*&*&*&*&1\\ 0&0&0&*&*&*&1&1&1 \end{pmatrix} \qquad \begin{pmatrix} 0&0&2&*&*&*&0&2&2\\ *&*&*&*&*&*&*&*&2\\ *&*&*&*&*&*&*&*&2\\ 0&0&0&*&*&*&2&2&2 \end{pmatrix}$$

Fig. 3: The structure of the matrices in $\mathcal{S}^{(3,3)}_{4\times 9,4}$ considering $P_1 = \{0\}$ and $P_2 = \{1, 2\}$

- for $i = 2, \dots, m-1$, $A_i = w_i\beta = v_i$, where v_i is a q-ary string of length n avoiding all the patterns $0^k, 1^k, 2^k, \dots, (p-1)^k$;
- $A_m = \alpha^k v_m \beta^k$, where v_m is a q-ary string of length $n-2k$ avoiding all the patterns $0^k, 1^k, 2^k, \dots, (p-1)^k$.

  (With A_1, A_i and A_m we denote the first, the i-th and the m-th row of the array $A \in S^{(k,p)}_{m\times n,q}(\alpha, \beta)$.)

Note that, if $p = 2$, then $\mathcal{S}^{(k,p)}_{m\times n,q} = \mathcal{S}^{(k)}_{m\times n,q}$ given in Definition 1. In Figure 3, we illustrate the structure of the matrices in $\mathcal{S}^{(3,3)}_{4\times 9,4}$ with $P_1 = \{0\}$ and $P_2 = \{1, 2\}$, pointing out that the entries marked with $*$ must avoid $000, 111$ and 222.

With the same argument used in Proposition 1, it is not difficult to prove the following proposition.

Proposition 2 *The set $\mathcal{S}^{(k,p)}_{m\times n,q} \subset \mathcal{M}_{m\times n,q}$ is non-overlapping, for each k, p, m, n with $2 \le k \le \lfloor \frac{n}{2} \rfloor$, $2 \le p \le q$ and $m \ge 2$.*

3 The enumeration of $\mathcal{S}^{(k,p)}_{m\times n,q}$

In this section we are going to enumerate the set $\mathcal{S}^{(k,p)}_{m\times n,q}$. First of all we can observe that, given $P = P_1 \cup P_2 = \{0, 1, 2, \dots, j-1\} \cup \{j, j+1, \dots, p-1\}$, then, for each $(\alpha, \beta) \in P_1 \times P_2$,

$$|\mathcal{S}^{(k,p)}_{m\times n,q}| = \sum_{(\alpha,\beta)} |\mathcal{S}^{(k,p)}_{m\times n,q}(\alpha, \beta)| \ ,$$

Since, $|\mathcal{S}^{(k,p)}_{m\times n,q}(\alpha, \beta)| = |\mathcal{S}^{(k,p)}_{m\times n,q}(\alpha', \beta')|$ for any $(\alpha, \beta), (\alpha', \beta') \in P_1 \times P_2$, then, fixed j with $1 \le j \le p-1$, we have

$$|\mathcal{S}^{(k,p)}_{m\times n,q}| = j(p-j)|\mathcal{S}^{(k,p)}_{m\times n,q}(\alpha, \beta)| \tag{1}$$

for some $(\alpha, \beta) \in P_1 \times P_2$. In the above formula $j(p-j)$ counts the number of pairs (α, β) which is of course maximum for $j = \lfloor \frac{p}{2} \rfloor$.

In order to evaluate $|\mathcal{S}^{(k,p)}_{m\times n,q}(\alpha, \beta)|$ it is easy to realize that it depends on the number of rows satisfying the constraints of Definition 2. We denote by $R^{(k,p)}_{n,q}$ the set of q-ary strings of length n starting with α, ending with β and avoiding all the patterns $0^k, 1^k, 2^k, \dots, (p-1)^k$. Let $Z^{(k,p)}_{n,q}$ be the set of q-ary strings of length n ending with β and avoiding all the patterns $0^k, 1^k, 2^k, \dots, (p-1)^k$. Moreover, let $B^{(k,p)}_{n,q}$ the set of q-ary strings of length n avoiding $0^k, 1^k, 2^k, \dots, (p-1)^k$. We indicate with $r_{n,q}^{(k,p)}$, $z_{n,q}^{(k,p)}$ and $b_{n,q}^{(k,p)}$ the cardinality of $R_{n,q}^{(k,p)}$, $Z_{n,q}^{(k,p)}$ and $B_{n,q}^{(k,p)}$, respectively. It is straightforward that

$$|\mathcal{S}_{m\times n,q}^{(k,p)}(\alpha,\beta)| = r_{n-2k+2,q}^{(k,p)} \cdot \left(z_{n,q}^{(k,p)}\right)^{m-2} \cdot b_{n-2k,q}^{(k,p)} \qquad (2)$$

where, referring to Definition 2, the term $r_{n-2k+2,q}^{(k,p)}$ counts the number of strings v_1, the terms $z_{n,q}^{(k,p)}$ counts the number of strings v_i for $i = 2, 3, \ldots, m-1$, and $b_{n-2k,q}^{(k,p)}$ is the number of strings v_m.

3.1 The sequence $b_{n,q}^{(k,p)}$

Now we consider a possible recursive relation for $b_{n,q}^{(k,p)}$ by means of a recursive construction of $B_{n,q}^{(k,p)}$. We first observe that $B_{0,q}^{(k,p)} = \{\lambda\}$, where λ is the empty string, and if $n < k$, then $B_{n,q}^{(k,p)}$ is formed by all the q-ary strings of length n. Therefore $b_{n,q}^{(k,p)} = q^n$ for $0 \le n < k$.

Denote by $B_{n,q}^{(k,p)}(u^i)$ the set of the strings in $B_{n,q}^{(k,p)}$ ending with exactly i equal symbols $u \in \Sigma$ and denote by $B_{n,q}^{(k,p)}(u) = \bigcup_{i=1}^{k-1} B_{n,q}^{(k,p)}(u^i)$ the set of strings in $B_{n,q}^{(k,p)}$ ending with at least a symbol u (note that $B_{n,q}^{(k,p)}(u)$ and $B_{n,q}^{(k,p)}(u^1)$ are different sets, in particular $B_{n,q}^{(k,p)}(u) \supset B_{n,q}^{(k,p)}(u^1)$).

Clearly,

$$B_{n,q}^{(k,p)} = \bigcup_{u=0}^{q-1} B_{n,q}^{(k,p)}(u).$$

In particular, $B_{n,q}^{(k,p)}$ can be partitioned in the strings ending with at most $k-1$ equal consecutive symbols of Σ and the strings ending with at least k equal consecutive symbols in $\Sigma \setminus P$. The first set can be generated from the strings in $B_{n-i,q}^{(k,p)}(u)$ followed by the suffix v^i, with $0 < i < k$, for each $v \in \Sigma$ with $u \neq v$, ranging u in Σ. The second set can be generated from the strings in $B_{n-k,q}^{(k,p)}(u)$ followed by the suffix v^k, for each $v \in \Sigma\setminus\{0, 1, 2, \ldots, p-1\}$, ranging u in Σ. This recursive construction assures the avoidance of the given forbidden patterns. Therefore,

$$b_{n,q}^{(k,p)} = \begin{cases} q^n & \text{if } 0 \le n \le k-1 \\ (q-1)\displaystyle\sum_{i=1}^{k-1} b_{n-i,q}^{(k,p)} + (q-p)b_{n-k,q}^{(k,p)} & \text{if } n \ge k. \end{cases} \qquad (3)$$

3.2 The sequences $z_{n,q}^{(k,p)}$

If $1 \le n < k$, then it easily seen that $z_{n,q}^{(k,p)} = q^{n-1}$. For $n \ge k$, the strings in $Z_{n,q}^{(k,p)}$ can be enumerated in terms of the strings in $B_{n,q}^{(k,p)}$. First of all, we note that

$Z^{(k,p)}_{n,q} = B^{(k,p)}_{n,q}(\beta)$. Then, the set $Z^{(k,p)}_{n,q}$ can be partitioned in $k-1$ subsets which are $B^{(k,p)}_{n,q}(\beta^i)$, with $i = 1, 2, \ldots, k-1$. The strings in $B^{(k,p)}_{n,q}(\beta^i)$ can be obtained appending the suffix β^i to all the strings contained in the set $B^{(k,p)}_{n-i,q}$ except the ones contained in $B^{(k,p)}_{n-i,q}(\beta)$ which coincides with $Z^{(k,p)}_{n-i,q}$.

In other words, for $i = 1, 2, \ldots, k-1$, it is $|B^{(k,p)}_{n,q}(\beta^i)| = b^{(k,p)}_{n-i,q} - z^{(k,p)}_{n-i,q}$. Since $Z^{(k,p)}_{n,q} = \bigcup_{i=1}^{k-1} B^{(k,p)}_{n,q}(\beta^i)$, and recalling the initial conditions, we have:

$$z^{(k,p)}_{n,q} = \begin{cases} 1 & \text{if } n = 0 \\ q^{n-1} & \text{if } 1 \leq n \leq k-1 \\ \displaystyle\sum_{i=1}^{k-1} \left(b^{(k,p)}_{n-i,q} - z^{(k,p)}_{n-i,q} \right) & \text{if } n \geq k. \end{cases} \tag{4}$$

Actually, it is possible to obtain an expression for $z^{(k,p)}_{n,q}$ involving only the two terms $b^{(k,p)}_{n,q}$ and $b^{(k,p)}_{n-1,q}$. More precisely, it can be observed that, for $n \geq 1$, the strings in $B^{(k,p)}_{n,q}$ can be partitioned in the strings of $Z^{(k,p)}_{n,q}$ ending with β for each $\beta \in P$ and the strings obtained by appending each symbols of $\Sigma \setminus P$ to each string in $B^{(k,p)}_{n-1,q}$. In other words $b^{(k,p)}_{n,q} = p\, z^{(k,p)}_{n,q} + (q-p) b^{(k,p)}_{n-1,q}$, hence

$$z^{(k,p)}_{n,q} = \frac{b^{(k,p)}_{n,q} - (q-p) b^{(k,p)}_{n-1,q}}{p}, \quad \text{for } n \geq 1\ . \tag{5}$$

3.3 The sequence $r^{(k,p)}_{n,q}$

Similarly to the above section it is easily seen that $R^{(k,p)}_{n,q} = Z^{(k,p)}_{n,q}(\alpha)$, where with $Z^{(k,p)}_{n,q}(\alpha)$ we denote the set containing the strings in $Z^{(k,p)}_{n,q}$ starting with α. Moreover, the set $R^{(k,p)}_{n,q}$ can be partitioned in $k-1$ subsets, say $Z^{(k,p)}_{n,q}(\alpha^i)$, with $i = 1, 2, \ldots, k-1$. The strings in $Z^{(k,p)}_{n,q}(\alpha^i)$ can be obtained appending the prefix α^i to all the strings contained in the set $Z^{(k,p)}_{n-i,q}$ except the ones contained in $Z^{(k,p)}_{n-i,q}(\alpha)$ which coincides with $R^{(k,p)}_{n-i,q}$. In other words, for $i = 1, 2, \ldots, k-1$, it is $|Z^{(k,p)}_{n,q}(\alpha^i)| = z^{(k,p)}_{n-i,q} - r^{(k,p)}_{n-i,q}$. Since $R^{(k,p)}_{n,q} = \bigcup_{i=1}^{k-1} Z^{(k,p)}_{n,q}(\alpha^i)$, and considering the easy initial conditions, we have:

$$r^{(k,p)}_{n,q} = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n = 1 \\ q^{n-2} & \text{if } 2 \leq n \leq k \\ \displaystyle\sum_{i=1}^{k-1} \left(z^{(k,p)}_{n-i,q} - r^{(k,p)}_{n-i,q} \right) & \text{if } n \geq k+1. \end{cases} \tag{6}$$

Also in this case it is possible to provide a formula for $r_{n,q}^{(k,p)}$ similar to (5). Nevertheless, here an additional term depending on k and n is needed. More precisely, we have $z_{n,q}^{(k,p)} = p\ r_{n,q}^{(k,p)} + (q-p)z_{n-1,q}^{(k,p)} + d_n^k$, where

$$d_n^{(k)} = \begin{cases} 1 \text{ if } (n \bmod k) = 0 \\ -1 \text{ if } (n \bmod k) = 1 \\ 0 \text{ if } (n \bmod k) \geq 2\ . \end{cases} \tag{7}$$

leading to the following proposition.

Proposition 3 *The sequence* $r_{n,q}^{(k,p)}$ *can be expressed by*

$$r_{n,q}^{(k,p)} = \frac{z_{n,q}^{(k,p)} - (q-p)z_{n-1,q}^{(k,p)} + d_n^{(k)}}{p} \quad \textit{for } n \geq 2\ . \tag{8}$$

Proof. We can proceed by induction on n. If $n = 2$, from expression (6) we have $r_{2,q}^{(k,p)} = 1$ which is the same value we obtain from (8) by replacing $z_{2,q}^{(k,p)}$, $z_{1,q}^{(k,p)}$ and $d_2^{(k)}$ with the corresponding value calculated from (4) and (7). Note that for $k = 2$, it is $z_{2,q}^{(2,p)} = b_{1,q}^{(2,p)} - z_{1,q}^{(2,p)} = q - 1$ (using (3)) and $d_2^{(2)} = 1$, while for $k \geq 3$, it is $z_{2,q}^{(k,p)} = q$, $a_{1,q}^{(k,p)} = 1$ and $d_2^{(k)} = 0$.

For $n > 2$ we distinguish the following three cases:

- for $2 < n < k$, from (6) we have $r_{n,q}^{(k,p)} = q^{n-2}$ which can be easily obtained from (8) using the expressions (4) and (7) for $z_{n,q}^{(k,p)}$ and $d_n^{(k)}$;
- for $n = k$, it is $r_{k,q}^{(k,p)} = q^{k-2}$ and, from (8),

$$\begin{aligned} r_{k,q}^{(k,p)} &= \frac{\sum_{i=1}^{k-1} \left(b_{k-i,q}^{(k,p)} - z_{k-i,q}^{(k,p)} \right) - (q-p)q^{k-2} + 1}{p} = \dots \\ &= \frac{q^{k-1} - 1 - (q-p)q^{k-2} + 1}{p} = q^{k-2} \quad ; \end{aligned}$$

- for $n > k$, we suppose that $r_{s,q}^{(k,p)} = \frac{z_{s,q}^{(k,p)} - (q-p)z_{s-1,q}^{(k,p)} + d_s^{(k)}}{p}$ for each $s < n$. Evaluating $r_{n,q}^{(k,p)}$ by means (6), using the inductive hypothesis and expression (4), we get:

$$
\begin{aligned}
r_{n,q}^{(k,p)} &= \sum_{i=1}^{k-1}\left(z_{n-i,q}^{(k,p)} - r_{n-i,q}^{(k,p)}\right) \\
&= \sum_{i=1}^{k-1}\left(\frac{b_{n-i,q}^{(k,p)} - (q-p)b_{n-i-1,q}^{(k,p)}}{p} - \frac{z_{n-i,q}^{(k,p)} - (q-p)z_{n-i-1,q}^{(k,p)} + d_{n-i}^{(k)}}{p}\right) \\
&= \tfrac{1}{p}\sum_{i=1}^{k-1}\left(b_{n-i,q}^{(k,p)} - z_{n-i,q}^{(k,p)}\right) - \frac{(q-p)}{p}\sum_{i=1}^{k-1}\left(b_{n-i-1,q}^{(k,p)} - z_{n-i-1,q}^{(k,p)}\right) - \frac{1}{p}\sum_{i=1}^{k-1} d_{n-i}^{(k)} \\
&= \frac{z_{n,q}^{(k,p)} - (q-p)z_{n-1,q}^{(k,p)} - \sum_{i=1}^{k-1} d_{n-i}^{(k)}}{p}\ .
\end{aligned}
$$

Since $\sum_{i=0}^{k-1} d_{n-i}^{(k)} = 0$, the term $\sum_{i=1}^{k-1} d_{n-i}^{(k)}$ is equal to d_n. Therefore

$$
r_{n,q}^{(k,p)} = \frac{z_{n,q}^{(k,p)} - (q-p)z_{n-1,q}^{(k,p)} + d_n^{(k)}}{p} \quad \text{for } n \geq 2\ ,
$$

as required.

■

3.4 Generating functions

The generating functions $b_q^{(k,p)}(x)$, $z_q^{(k,p)}(x)$ and $r_q^{(k,p)}(x)$ for the sequences (3), (4) and (6), respectively, can be easily determined by using the general approach presented in [22] adapted to our constraints. Here, we are going to directly determine them from the recurrences in the previous section. From (3), we obtain:

$$
\begin{aligned}
b_q^{(k,p)}(x) &= \sum_{n\geq 0} b_{n,q}^{(k,p)} x^n \\
&= b_{0,q}^{(k,p)} + b_{1,q}^{(k,p)}x + b_{2,q}^{(k,p)}x^2 + \ldots + b_{k-1,q}^{(k,p)}x^{k-1} + \sum_{n\geq k} b_{n,q}^{(k,p)} x^n \\
&= \sum_{n=0}^{k-1} q^n x^n + \sum_{n\geq k}\left((q-1)\sum_{i=1}^{k-1} b_{n-i,q}^{(k,p)} + (q-p)b_{n-k,q}^{(k,p)}\right)x^n \\
&= \sum_{n=0}^{k-1} q^n x^n + (q-1)\sum_{i=1}^{k-1} x^i \sum_{n\geq k} b_{n-i,q}^{(k,p)} x^{n-i} + (q-p)x^k \sum_{n\geq k} b_{n-k,q}^{(k,p)} x^{n-k} \\
&= \sum_{n=0}^{k-1} q^n x^n + (q-1)\sum_{i=1}^{k-1} x^i\left(b_q^{(k,p)}(x) - \sum_{j=0}^{k-i-1} q^j x^j\right) + (q-p)x^k b_q^{(k,p)}(x).
\end{aligned}
$$

Since $(q-1)\sum_{i=1}^{k-1}\sum_{j=i}^{k-1} q^{j-i}x^j = \sum_{j=1}^{k-1} q^j x^j - \sum_{i=1}^{k-1} x^i$ then we deduce that

$$b_q^{(k,p)}(x) = \sum_{n=0}^{k-1} q^n x^n + (q-1)b_q^{(k,p)}(x)\sum_{i=1}^{k-1} x^i - \sum_{j=1}^{k-1} q^j x^j + \sum_{i=1}^{k-1} x^i + (q-p)x^k b_q^{(k,p)}(x).$$

Hence,

$$b_q^{(k,p)}(x) = \frac{\sum_{i=0}^{k-1} x^i}{1-(q-1)\sum_{i=1}^{k-1} x^i - (q-p)x^k}. \tag{9}$$

From recurrences (4) and (6) the generating functions $z_q^{(k,p)}(x)$ and $r_q^{(k,p)}(x)$ can be easily obtained:

$$z_q^{(k,p)}(x) = 1 + \frac{b_q^{(k,p)}(x)\sum_{i=1}^{k-1} x^i}{\sum_{i=0}^{k-1} x^i}. \tag{10}$$

$$r_q^{(k,p)}(x) = 1 + \frac{\left(z_q^{(k,p)}(x)-1\right)\sum_{i=1}^{k-1} x^i}{\sum_{i=0}^{k-1} x^i}. \tag{11}$$

Moreover, it is possible to observe that the generating function for the sequence $\left\{|\mathcal{S}_{m\times n,q}^{(k,p)}|\right\}_{n\geq 2k}$, for any fixed m, p, q and k, using relations (1), (2) and the above expressions for the generating functions, is rational due to the fact that the Hadamard product of rational generating functions is rational (see [25]).

3.5 Asymptotic behavior

First, we estimate the asymptotic behaviour of the n-th coefficient $b_{n,q}^{(k,p)}$ of $b_q^{(k,p)}(x)$. Following [20] we say that the sequence $b_{n,q}^{(k,p)}$ is of *exponential order* C^n (abbreviated by $b_{n,q}^{(k,p)} \bowtie C^n$) if, for any $\varepsilon > 0$,

- $|b_{n,q}^{(k,p)}| >_{i.o.} (C-\varepsilon)^n$; that is to say, $|b_{n,q}^{(k,p)}|$ exceeds $(C-\varepsilon)^n$ *infinitely often* (for infinitely many values of n);
- $|b_{n,q}^{(k,p)}| <_{a.e.} (C+\varepsilon)^n$; that is to say, $|b_{n,q}^{(k,p)}|$ is dominated by $(C+\varepsilon)^n$ *almost everywhere* (except for possibly finitely many values of n).

The following theorem holds (Exponential Growth Formula, [20]):

Theorem 1 *If $f(x)$ is analytic at 0 and x_0 is the modulus of a singularity nearest to the origin, then the coefficient $f_n = [x^n]f(x)$ satisfies*

$$f_n \bowtie \left(\frac{1}{x_0}\right)^n .$$

Since $b_q^{(k,p)}(x) = \frac{h(x)}{g(x)}$ is a rational function, the modulus $x_0 = x_0(k,p,q)$ of a singularity of $b_q^{(k,p)}(x)$ nearest to the origin is the modulus of a zero of $g(x)$ nearest to the origin. We have the following proposition:

Proposition 4 *The polynomial $g(x)$ has a unique zero in $A = \left\{x \in \mathbb{C} : |x| < \frac{1}{q-1}\right\}$, for $q \geq 3$.*

Proof. We consider the polynomial

$$G(x) = (1-x)g(x) = x^k(q-1) - qx + 1 - (1-x)(q-p)x^k.$$

The zeros of $G(x)$ are the zeros of $g(x)$ and $\bar{x} = 1$. Clearly, if x_0 is the modulus of the zero nearest to the origin of $G(x)$ and $x_0 < 1$, then x_0 is also the modulus of the zero nearest to the origin of $g(x)$.

Since $G(1/q) = \frac{p(q-1)}{q^{k+1}} > 0$ and $G\left(\frac{1}{q-1}\right) = \frac{1-(q-1)^{k-2}}{(q-1)^{k-1}} - \frac{(q-2)(q-p)}{(q-1)^{k+1}} < 0$, for $k, q \geq 3$, then $G(x)$ has at least a real zero α, with $\frac{1}{q} < \alpha < \frac{1}{q-1}$. We will prove that α is the unique zero of $G(x)$ inside A. Therefore, $x_0 = \alpha$ is the modulus of the singularity nearest to the origin of $b_q^{(k,p)}(x)$. This proof is carried on by means Rouché's Theorem (see for example [30]). To this aim we pose $G(x) = F_1(x) + F_2(x)$ where

$$F_1(x) = 1 - qx$$

and

$$F_2(x) = x^k(q-1) - (1-x)(q-p)x^k.$$

It is easily seen that, for $x \in \partial A = \left\{x \in \mathbb{C} : |x| = \frac{1}{q-1}\right\}$,

$$|F_1(x)| \geq \frac{1}{q-1}$$

and

$$|F_2(x)| \leq (p-1)|x|^k + (q-p)|x|^{k+1} = \frac{pq - 2p + 1}{(q-1)^{k+1}}.$$

Since $\frac{pq-2p+1}{(q-1)^{k+1}} < \frac{1}{q-1}$ for $k, q \geq 3$, we have

$$|F_2(x)|_{x \in \partial A} < |F_1(x)|_{x \in \partial A}.$$

Clearly, $F_1(x)$ has the unique zero $\frac{1}{q} \in A$. Then, for Rouché's Theorem, we have that $G(x) = F_1(x) + F_2(x)$ has a unique zero $x_0 = \alpha \in A$, which is also the modulus of the zero of $g(x)$ nearest to the origin.

■

By Theorem 1 the coefficient $b_{n,q}^{(k,p)}$ is of exponential order $\left(\frac{1}{x_0(k,p,q)}\right)^n$. Being $b_q^{(k,p)}(x) = \frac{h(x)}{g(x)}$, we observe that, from (10), $z_q^{(k,p)}(x) = 1 + \frac{h(x)-1}{g(x)}$ has the same denominator of $b_q^{(k,p)}(x)$ so that x_0 is the smallest singularity of $z_q^{(k,p)}(x)$. Moreover, from (11), it is $r_q^{(k,p)}(x) = 1 + \frac{(h(x)-1)^2}{g(x)h(x)}$. Since all the zeros of $h(x)$ have modulus equal to 1 and $x_0 < 1$, the smallest singularity of $r_q^{(k,p)}(x)$ is x_0 again.

Summarizing, we have the following proposition:

Proposition 5 *For $q \geq 3$,*

$$\left(\frac{1}{x_0(k,p,q)} - \varepsilon\right)^n <_{i.o.} r_{n,q}^{(k,p)} < z_{n,q}^{(k,p)} < b_{n,q}^{(k,p)} <_{a.e.} \left(\frac{1}{x_0(k,p,q)} + \varepsilon\right)^n . \quad (12)$$

Therefore $r_{n,q}^{(k,p)}$, $z_{n,q}^{(k,p)}$ and $b_{n,q}^{(k,p)}$ are of exponential order $\left(\frac{1}{x_0(k,p,q)}\right)^n$.

For what the asymptotic behavior of $|\mathcal{S}_{m\times n,q}^{(k,p)}|$ is concerned, we recall that, according to (1), (2) and considering $j = \left\lfloor \frac{p}{2} \right\rfloor$, the size of $\mathcal{S}_{m\times n,q}^{(k,p)}$ is given by

$$|\mathcal{S}_{m\times n,q}^{(k,p)}| = \left\lfloor \frac{p}{2} \right\rfloor \left\lceil \frac{p}{2} \right\rceil \left(r_{n-2k+2,q}^{(k,p)} \cdot \left(z_{n,q}^{(k,p)}\right)^{m-2} \cdot b_{n-2k,q}^{(k,p)} \right) . \quad (13)$$

From Proposition 5, it immediately follows:

Proposition 6 *For $q \geq 3$,*

$$\left\lfloor \frac{p}{2} \right\rfloor \left\lceil \frac{p}{2} \right\rceil \left(\frac{1}{x_0(k,p,q)} - \varepsilon\right)^{nm-4k+2} <_{i.o.} |\mathcal{S}_{m\times n,q}^{(k,p)}|$$

and

$$|\mathcal{S}_{m\times n,q}^{(k,p)}| <_{a.e.} \left\lfloor \frac{p}{2} \right\rfloor \left\lceil \frac{p}{2} \right\rceil \left(\frac{1}{x_0(k,p,q)} + \varepsilon\right)^{nm-4k+2} .$$

The above proposition shows that the asymptotic behavior of $|\mathcal{S}_{m\times n,q}^{(k,p)}|$ is the same when $n \to \infty$ or $m \to \infty$.

The exact expression for $x_0(k,p,q)$ can not be easily found. Nevertheless we provide a lower and an upper bound x_1 and x_2 for x_0 depending on the parameters k, p and q.

A first fact is that $\frac{1}{q} < x_0 < \frac{1}{q-1}$. Using classical approximation tools it is possible to obtain better bounds for x_0. In particular, for a left bound x_1, we observe that, being $g'(x) < 0$, $g''(x) < 0$ for $\frac{1}{q} < x < \frac{1}{q-1}$, and

$$g\left(\frac{1}{q}\right) = \frac{p}{q^k} > 0, \qquad g\left(\frac{1}{q-1}\right) = -\frac{(q-1)^k - p(q-2) - 1}{(q-1)^k(q-2)} < 0 ,$$

then the straight line passing through $\left(\frac{1}{q}, g\left(\frac{1}{q}\right)\right)$ and $\left(\frac{1}{q-1}, g\left(\frac{1}{q-1}\right)\right)$ crosses the x-axis in x_1 such that $\frac{1}{q} < x_1 < x_0$. It is not difficult to obtain that

$$\begin{aligned} x_1 &= \frac{1}{q} + \frac{p}{q(q-1)\left(p - g\left(\frac{1}{q-1}\right)q^k\right)} \\ &= \frac{1}{q} + \frac{p(q-1)^{k-1}(q-2)}{q(p(q-1)^k(q-2) + q^k(q-1)^k - pq^k(q-2) - q^k)} . \end{aligned}$$

A better right bound x_2 for x_0 can be found by means of Newton's method for finding successively better approximations to the root of $g(x)$ with initial point $\frac{1}{q}$. The above consideration about the derivatives of $g(x)$ assures that $x_2 > x_0$. Its exact expression is:

$$x_2 = \frac{1}{q} - \frac{g\left(\frac{1}{q}\right)}{g'\left(\frac{1}{q}\right)} = \frac{1}{q} + \frac{p(q-1)}{q(q^{k+1} - kp(q-1) - q)} \ .$$

It is easy to check that $x_2 < \frac{1}{q-1}$, therefore

$$\frac{1}{q} < x_1 < x_0 < x_2 < \frac{1}{q-1} \ ,$$

hence

$$q - 1 < \frac{1}{x_2} < \frac{1}{x_0} < \frac{1}{x_1} < q \ , \tag{14}$$

with

$$\frac{1}{x_1} = q - \frac{p(q-1)^{k-1}(q-2)}{p(q-1)^{k-1}(q-2) + q^{k-1}(q-1)^k - pq^{k-1}(q-2) - q^{k-1}} \tag{15}$$

and

$$\frac{1}{x_2} = q - \frac{qp(q-1)}{q(q^k-1) - p(q-1)(k-1)} \ . \tag{16}$$

For the sake of completeness we also consider the case $q = 2$. Recalling that $2 \le p \le q$, from (9) it is

$$b_2^{(k,2)}(x) = \frac{\sum_{i=0}^{k-1} x^i}{1 - \sum_{i=1}^{k-1} x^i} \ .$$

The modulus $x_0 = x_0(k)$ of a singularity of $b_2^{(k,2)}(x)$ nearest to the origin is the modulus of a zero of the denominator nearest to the origin. It can be shown that $x_0 = \frac{1}{y_0}$ where y_0 is the positive real root of the polynomial

$$y(x) = y^{k-1} - y^{k-2} - \ldots - y - 1 \ .$$

The equation $y(x) = 0$ is studied in [31] where it can be found that $y_0 = 2(1-\delta_k)$ with $\delta_k = \sum_{i\ge1} \binom{ki-2}{i-1} \frac{1}{i2^{ki}}$. From Theorem 1 we deduce

$$b_{n,2}^{(k,2)} \bowtie (2(1-\delta_k))^n \ .$$

With an argument similar to the one used for the singularities nearest to the origin of $z_q^{(k,p)}(x)$ and $r_q^{(k,p)}(x)$, we deduce that the modulus of the singularities nearest to the origin of $z_2^{(k,2)}(x)$ and $r_2^{(k,2)}(x)$ are $1/y_0$ again. Hence, from Theorem 1, we have:

Proposition 7 *For $q = 2$,*

$$(2(1-\delta_k)-\varepsilon)^n <_{i.o.} r_{n,2}^{(k,2)} < z_{n,2}^{(k,2)} < b_{n,2}^{(k,2)} <_{a.e.} (2(1-\delta_k)+\varepsilon)^n \ . \tag{17}$$

Therefore $r_{n,2}^{(k,2)}$, $z_{n,2}^{(k,2)}$ and $b_{n,2}^{(k,2)}$ are of exponential order $(2(1-\delta_k))^n$.

The asymptotic behavior of $|\mathcal{S}_{m\times n,2}^{(k,p)}|$ is easily obtained from the previous proposition and formula (13).

4 On the maximal size of $\mathcal{S}_{m\times n,q}^{(k,p)}$

In this section we provide some considerations about the maximal size of $\mathcal{S}_{m\times n,q}^{(k,p)}$ taking into account its asymptotic behavior. More precisely, we are interested in the values of k and p giving the maximal value of (13) when n is sufficiently large, once q is fixed. A first step in this direction is the analysis of the sequences involved in (13). We start from $z_{n,q}^{(k,p)}$. It is already known (see Proposition 5) that

$$z_{n,q}^{(k,p)} \bowtie \left(\frac{1}{x_0(k,p,q)}\right)^n$$

and

$$\frac{1}{x_2} < \frac{1}{x_0(k,p,q)} < \frac{1}{x_1}$$

where the values of the bounds can be found in (15) and (16). We observe that $\frac{1}{x_1} = q - \frac{p(q-1)^{k-1}(q-2)}{p(q-1)^{k-1}(q-2)+q^{k-1}(q-1)^k - pq^{k-1}(q-2)-q^{k-1}} < q - \frac{p(q-1)^{k-1}(q-2)}{q^{k-1}(q-1)^k - q^{k-1}}$ hence

$$q - \frac{qp(q-1)}{q(q^k-1)-p(q-1)(k-1)} < \frac{1}{x_0(k,p,q)} < q - \frac{p(q-1)^{k-1}(q-2)}{q^{k-1}(q-1)^k - q^{k-1}} \ . \tag{18}$$

If k increases it is easily seen that the same happens for $1/x_1$ and $1/x_2$, so that also $1/x_0$ increases. Moreover, if p decreases, then $1/x_1$ and $1/x_2$, and so $1/x_0$, increase. Therefore the number $z_{n,q}^{(k,p)}$ of strings avoiding the patterns 0^k, 1^k, ..., $(p-1)^k$ rises when the length k of the patterns increases and the number p of the forbidden patterns decreases. Then the maximal size of $z_{n,q}^{(k,p)}$, which counts the number of strings in each row v_i for $i = 2, 3, \ldots, m-1$ of a given matrix in $\mathcal{S}_{m\times n,q}^{(k,p)}$, is obtained when k is the greatest possible value ($k = \lfloor \frac{n}{2} \rfloor$) and p is the smallest one ($p = 2$).

The above discussion can be used as a formal argument to justify the intuitive fact that a set of strings avoiding a certain set of p patterns with a certain length k increases in its cardinality if the number of patterns decreases and their length increases.

Thanks to Proposition 5 the asymptotic behavior of $z_{n,q}^{(k,p)}$ is similar to the one of $b_{n,q}^{(k,p)}$ and $r_{n,q}^{(k,p)}$. Notice that for the estimation of formula (13) we have to recall that the strings in the first and the last row of a matrix in $\mathcal{S}_{m\times n,q}^{(k,p)}$ have length $n-2k+2$ and $n-2k$, respectively, instead of n. Therefore we have

$$r^{(k,p)}_{n-2k+2,q} \bowtie \left(\frac{1}{x_0(k,p,q)}\right)^{n-2k+2}$$

and

$$b^{(k,p)}_{n-2k,q} \bowtie \left(\frac{1}{x_0(k,p,q)}\right)^{n-2k}$$

whit the same bounds (18) for $1/x_0(k,p,q)$. Since the length k appears also in the exponent as well as in $1/x_2$ and $1/x_1$, the coefficients $r^{(k,p)}_{n-2k+2,q}$ and $b^{(k,p)}_{n-2k,q}$ do not have the same behavior of $z^{(k,p)}_{n,q}$ when k increases. More precisely, experimental results show a unimodal behavior of $r^{(k,p)}_{n-2k+2,q}$ and $b^{(k,p)}_{n-2k,q}$ depending on k. Another difference with respect to $z^{(k,p)}_{n,q}$ can be observed when the cardinality of the alphabet grows. In particular, since $\lim_{q\to\infty}(1/x_1)/q = \lim_{q\to\infty}(1/x_2)/q = 1$, there exists a q_l such that $1/x_1 \simeq q$ and $1/x_2 \simeq q$, if $q > q_l$. In other words, $1/x_1$ and $1/x_2$ can be arbitrarily close to q so that the coefficients $r^{(k,p)}_{n-2k+2,q}$ and $b^{(k,p)}_{n-2k,q}$ can be approximated by q^{n-2k+2} and q^{n-2k}, respectively. Therefore, their greatest value is obtained for $k = 2$, as opposed to the value $k = \lfloor \frac{n}{2} \rfloor$ giving the maximum for $z^{(k,p)}_{n,q}$, independently of q.

As far as the parameter p is concerned, notice that the coefficients $r^{(k,p)}_{n-2k+2,q}$ and $b^{(k,p)}_{n-2k,q}$ have the same behavior as $z^{(k,p)}_{n,q}$ when p decreases, then they assume their maximum value when $p = 2$, as opposed to the value $p = q$ giving the maximum for the term $\lfloor \frac{p}{2} \rfloor \lceil \frac{p}{2} \rceil$ in (13).

From the above arguments it follows that, in general, it is not easy to find the maximum value of $\mathcal{S}^{(k,p)}_{m\times n,q}$ depending on k and p. Some numerical results of this fact are showed in Table 1.

5 Further developments

The present work is the generalization of the study started in [5] where we considered sets constituted by only square matrices and overlappings only along the direction of the main diagonal. Surely, these hypotheses lead to a very special case and they are very restrictive. Nevertheless, such sets of square matrices have at least two notable properties: the matrices can be listed in a 1-Gray code sense (by using the Gray code defined in [9] and the well-known generalization [19] of the Binary Reflected Gray Code [21]) and the sets are non-expandable. As far as the 1-Gray code (which is a list of the matrices of the set in a such way that two consecutive matrices differ only for one entry) is concerned, we observe that $\mathcal{S}^{(k,p)}_{m\times n,q}$ is not a 1-Gray code. Indeed, it is possible to prove that in general there does not exist a 1-Gray code for the strings in $B^{(k,2)}_{n,2}$ so that it does not exist for $\mathcal{S}^{(k,p)}_{m\times n,q}$. In order to show that the strings in $B^{(k,2)}_{n,2}$ can not be listed as desired, we note that a 1-Gray code exists if and only if it is possible to find an Hamiltonian path on the graph whose vertexes are the strings in $B^{(k,2)}_{n,2}$ and the edges connect

		$p = 2$	$p = 3$	$p = 4$
$\left\lvert \mathcal{S}^{(k,p)}_{3\times 8,4} \right\rvert$	$k = 2$	$\mathbf{1.21 \cdot 10^{8}}$	$9.36 \cdot 10^{7}$	$5.76 \cdot 10^{7}$
	$k = 3$	$3.05 \cdot 10^{6}$	$5.68 \cdot 10^{6}$	$1.05 \cdot 10^{7}$
	$k = 4$	$1.57 \cdot 10^{4}$	$3.11 \cdot 10^{4}$	$6.14 \cdot 10^{4}$
$\left\lvert \mathcal{S}^{(k,p)}_{8\times 8,4} \right\rvert$	$k = 2$	$1.18 \cdot 10^{27}$	$8.74 \cdot 10^{25}$	$2.88 \cdot 10^{24}$
	$k = 3$	$1.44 \cdot 10^{27}$	$1.87 \cdot 10^{27}$	$\mathbf{2.38 \cdot 10^{27}}$
	$k = 4$	$1.54 \cdot 10^{25}$	$2.84 \cdot 10^{25}$	$5.25 \cdot 10^{25}$
$\left\lvert \mathcal{S}^{(k,p)}_{20\times 8,4} \right\rvert$	$k = 2$	$4.46 \cdot 10^{72}$	$1.17 \cdot 10^{69}$	$3.45 \cdot 10^{64}$
	$k = 3$	$\mathbf{6.04 \cdot 10^{76}}$	$3.30 \cdot 10^{76}$	$1.68 \cdot 10^{76}$
	$k = 4$	$3.66 \cdot 10^{75}$	$5.78 \cdot 10^{75}$	$9.09 \cdot 10^{75}$
$\left\lvert \mathcal{S}^{(k,p)}_{60\times 8,4} \right\rvert$	$k = 2$	$3.67 \cdot 10^{224}$	$6.80 \cdot 10^{212}$	$1.35 \cdot 10^{198}$
	$k = 3$	$1.52 \cdot 10^{242}$	$4.69 \cdot 10^{240}$	$1.15 \cdot 10^{239}$
	$k = 4$	$\mathbf{3.05 \cdot 10^{243}}$	$2.84 \cdot 10^{243}$	$2.61 \cdot 10^{243}$

Table 1: Maximal values (bold character) of $\left\lvert \mathcal{S}^{(k,p)}_{m\times 8,4} \right\rvert$ as m grows

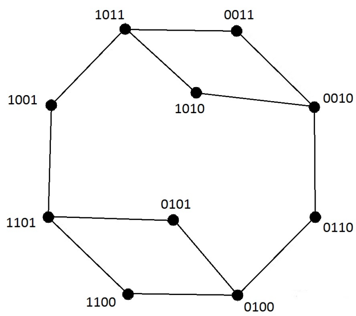

Fig. 4: The graph of $B^{(3,2)}_{4,2}$ non admitting an Hamiltonian path

any two strings differing for only one symbol of the alphabet. For example it does not exist an Hamiltonian path in the graph of $B^{(3,2)}_{4,2}$ (see Figure 4).

Nevertheless, a further development in this direction could be the investigation about the existence of a d-Gray code ($d \geq 2$), in the sense that any two consecutive matrices differ in at most d entries.

Moreover, recalling that a non-expandable set N is such that any other matrix $C \notin N$ can be overlapped with at least one matrix of N, it is not difficult to prove that the set $\mathcal{S}^{(k,p)}_{m\times n,q}$ considered in this work is not a non-expandable set. Indeed, considering $A \in \mathcal{S}^{(k,p)}_{m\times n,q}(\alpha,\beta)$ and $B \in \mathcal{S}^{(k+1,p)}_{m\times n,q}(\alpha,\beta)$, the last row of A can be overlapped with the first row of B. More precisely, if the last row of A is $A_m = \alpha^k u\beta^k$ and the first row of B is $B_1 = \alpha^k\beta u'\alpha\beta^k$, then A_m and B_1 can be overlapped posing $u = \beta u'\alpha$.

A further task could be the characterization of a set of non-overlapping matrices over a q-ary alphabet where the matrices have different dimensions. This would be a possible generalization in two dimensions of sets of variable length codes. Also, it could be fruitful to investigate on the possibility to make that set non-expandable.

6 Acknowledgements

This work has been partially supported by the PRIN project “Automi e linguaggi formali: aspetti matematici ed applicativi” and GNCS project “Strutture discrete con vincoli”.

References

1. Anselmo, M., Giammarresi, D., Madonia, M.: Two-Dimensional Rational Automata: A Bridge Unifying One- and Two-Dimensional Language Theory. Lecture Notes in Comput. Sci. 7741, 133–145 (2013)
2. Anselmo, M., Giammarresi, D., Madonia, M.: Structure and Measure of a Decidable Class of Two-dimensional Codes. Lecture Notes in Comput. Sci. 8977, 315–327 (2015)
3. Anselmo, M., Giammarresi, D., Madonia, M.: Unbordered Pictures: Properties and Construction. Lecture Notes in Comput. Sci. 9270, 45–57 (2015)
4. Bacher, A., Bernini, A., Ferrari, L., Gunby, B., Pinzani, R., West, J.: The Dyck pattern poset. Discrete Math. 321, 12–23 (2014)
5. Barcucci, E., Bernini, A., Bilotta, S., Pinzani, R.: Cross-bifix-free sets in two dimensions. Theoret. Comput. Sci. (2015). doi:10.1016/j.tcs.2015.08.032
6. Barcucci, E., Bernini, A., Bilotta, S., Pinzani, R.: Non-overlapping matrices. Theoret. Comput. Sci. (2016). doi:10.1016/j.tcs.2016.05.009
7. Barcucci, E., Bilotta, S., Pergola, E., Pinzani, R., Succi, J.: Cross-bifix-free generation via Motzkin paths. RAIRO Theor. Inform. Appl. 50, 81 – 91, (2016)
8. Bajic, D., Loncar-Turukalo, T.: A simple suboptimal construction of cross-bifix-free codes. Cryptogr. Commun. 6, 27–37 (2014)
9. Bernini, A., Bilotta, S., Pinzani, R., Sabri, A., Vajnovszki, V.: Prefix partitioned gray codes for particular cross-bifix-free sets. Cryptogr. Commun. 6, 359 – 369 (2014)
10. Bernini, A., Ferrari, L., Pinzani, R., West, J.: Pattern-avoiding Dyck paths. Discrete Math. Theoret. Comput. Sci. FPSAC 2013, 683–694 (2013)
11. Bernini, A., Ferrari, L., Steingrmsson, E.: The Möbius function of the consecutive pattern poset. Electron. J. Combin. 18, P146 (2011)
12. Bilotta, S., Grazzini, E., Pergola, E., Pinzani, R.: Avoiding cross-bifix-free binary words. Acta Inform. 50, 157–173 (2013)
13. Bilotta, S., Merlini, D., Pergola, E., Pinzani, R.: Pattern $1^{j+1}0^j$ avoiding binary words. Fund. Inform. 117, 35–55 (2012)
14. Bilotta, S., Pergola, E., Pinzani, R.: A new approach to cross-bifix-free sets. IEEE Trans. Inform. Theory 58, 4058–4063 (2012)
15. Björner, A.: The Möbius function of subword order. Invariant theory and tableaux (Minneapolis, MN, 1988), 118124, IMA Vol. Math. Appl., 19, Springer, New York, 1990.
16. Blackburn, S.: Non-overlapping codes. IEEE Trans. Inform. Theory 61, 4890–4894 (2015)
17. Burstein, A. Enumeration of words with forbidden patterns. PhD thesis, University of Pennsylvania, 1998.
18. Chee, Y. M., Kiah, H. M., Purkayastha, P., Wang, C.: Cross-bifix-free codes within a constant factor of optimality. IEEE Trans. Inform. Theory 59, 4668–4674 (2013)

19. Er, M. C.: On generating the N-ary reflected Gray code. IEEE Trans. Comput. 33, 739–741 (1984)
20. Flajolet, P., Sedgewick, R.: Analytic Combinaotrics Cambridge University Press, 2009
21. Gray, F.: Pulse Code Communication. U.S. Patent 2 632 058 (1953)
22. Guibas, L. J., Odlyzko, A. M.: String overlaps, pattern matching and nontransitive games. J. Combin. Theory Ser. A 30, 183 – 208 (1981)
23. Klazar, M.: On abab-free and abba-free sets partitions. European J. Combin. 17, 53–68 (1996)
24. Knuth, D.: The art of computer programming, Vol. 1. Addison Wesley, Boston, 1968
25. Lando, S. K.: Lecture on Generating Functions. American Mathematical Society, 2003
26. Otto, F., Mráz, F.: Extended two-way ordered restarting automata for picture languages. Lecture Notes in Comput. Sci. 8370, 541–552 (2014)
27. Pradella, M., Cherubini, A., Crespi-Reghizzi, S.: A unifying approach to picture grammars. Inform. and Comput. 209, 1246–1267 (2011)
28. Rowland, E.: Pattern avoidance in binary trees. J. Combin. Theory Ser. A 117, 741–758 (2010)
29. Sagan, B. E.: Pattern avoidance in set partitions Ars Combin. 117, 7996 (2010)
30. Spiegel, M. R.: Complex variables. McGraw-Hill, New York, 1964
31. Wolfram, D. A.: Solving generalized Fibonacci recurrences. Fibonacci Quart. 36, 129–145, (1998)