DAVENPORT-SCHINZEL THEORY OF MATRICES

Zoltán Füredi* and Péter Hajnal†

* Mathematical Institute of the Hungarian Academy of Sciences,
1364 Budapest, P. O. B. 127, Hungary

and

† Princeton University,
Department of Computer Science,
Princeton, NJ 08544

Abstract. Let C be a configuration of 1's. We define $f(n;C)$ to be the maximal number of 1's in a 0-1 matrix of size $n \times n$ not having C as a subconfiguration. We consider the problem of determining the order of $f(n;C)$ for several forbidden C's. Among others we prove that $f(n; \begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \end{pmatrix}) = \Theta(\alpha(n)n)$, where $\alpha(n)$ is the inverse of the Ackermann function.

1. Introduction

A *configuration*, $C = (c_{ij})$ $(1 \le i \le u,\ 1 \le j \le v)$, is a partial matrix with 1's and blanks at the entries. All the matrices we are going to work with will be $0-1$ matrices. We say that a matrix $M = (m_{ij})$ does have the configuration C if one can find u rows $i_1, i_2, \ldots, i_u, i_1 < \cdots < i_u$ and v columns $j_1, j_2, \ldots, j_v, j_1 < \cdots < j_v$ in M such that the corresponding submatrix contains C, i.e. $m_{i_\alpha, j_\beta} = 1$ whenever $c_{\alpha,\beta} = 1$. Let $f(n,m;C)$ denote the maximum number of 1's in an $n \times m$ matrix M not containing C. In the case of $n = m$ we write $f(n;C)$. One can allow several forbidden configurations, the corresponding threshold function is $f(n,m;\{C^1, \ldots, C^n\})$ or $f(n;\{C^1, \ldots, C^n\})$.

Our research is closely related to previous works in combinatorics.

First, let us mention Turán's theory in extremal graph theory. There the question is: Given a graph G, what is $T(n;G)$, the maximum number of edges of a graph with n vertices and not containing G as a subgraph? A special case is

* Research supported partly by the Hungarian National Science Foundation grant No. 1817. This paper was written while the first author visited Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139.

† Part of this work was carried out at M.I.T. This work was supported in part by the Air Force under Contract OSR-86-0076 and by DIMACS (Center for Discret Mathematics and Theoretical Computer Science), a National Science Foundation Science and Technology Center – NSF-STC88-09648. The second author is on leave of absence from Bolyai Institute, University of Szeged, Szeged, Hungary.

when we work in the universe of bipartite graphs. Our matrices can be considered as bipartite graphs. The important difference between Turán's theory and our question that in our case the vertices (the rows and columns) are ordered. This is a very important difference but in some special case the restriction on the order is insignificant. An example is the four cycle (complete bipartite graph between two color classes of size 2 each). Classical results in graph theory [KST], [ERS], [B] immediately give us the following theorem.

THEOREM 1.1. $f(n; \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}) = \Theta(n^{\frac{3}{2}})$ ∎.

We do not know exactly how these two problems are related, but the following facts are known. The Erdős-Stone-Simonovits theorem ([ESi], [ESt], for a survey see Bollobás' book [Bo]) says that the order of magnitude of $T(n;G)$ depends on the chromatic number of G, namely $\lim_{n\to\infty} \frac{T(n;G)}{\binom{n}{2}} = 1 - (\chi(G)-1)^{-1}$. This theorem gives sharp estimate on $T(n;G)$, except for bipartite G. For every bipartite graph B which is not a tree there are positive constants c_1 and c_2 (not depending on n) such that

$$\Omega(n^{1+c_1}) \leq T(n;B) \leq O(n^{2-c_2})$$

holds. If the graph is a tree F, then it is straightforward that $T(n;F) = \Theta(n)$. However we will see that our problem has completely different threshold functions. For a special matrix (such that the corresponding graph is a tree, hence it has linear Turán function) our threshold function turns out to be $\Theta(n \log n)$.

An other related question is raised by Davenport and Schinzel. A sequence $s = x_1 x_2 \ldots x_l$ is called a *Davenport-Schinzel sequence*, $s \in DS_k(n)$, if $x_i \neq x_{i+1}$, $x_i \in \{1, 2, \ldots, n\}$ and s does not contain a subsequence $x_{i_1} x_{i_2} \ldots x_{i_k}$ such that

$$x_{i_1} = x_{i_3} = \ldots = x_{i_{2t-1}} = \ldots \neq x_{i_2} = x_{i_4} = \ldots = x_{i_{2t}} = \ldots$$

$(i_1 < i_2 < \ldots < i_k)$. Let $ds_k(n)$ denote the maximum length of an $s \in DS_k(n)$. It is obvious that

$$ds_3(n) = n, \quad ds_4(n) = 2n - 1.$$

Szemerédi [Sz] proved that $ds_k(n) = O(n \log^*(n))$ for all fixed k while n tends to infinity. (Here, as usual, $\log^* n$ denotes the inverse of the function $p\colon \mathbf{N} \to \mathbf{N}$ with $p(1) = 2$, $p(n+1) = 2^{p(n)}$.) Recently, mainly due to the works of M. Sharir ([S], [HS], [GHS], [K]) it is known that the true order of the magnitude of $ds_k(n)$ for $k \geq 5$ is really superlinear, e.g. (Hart and Sharir [HS])

$$ds_5(n) = \Theta(n\alpha(n)),$$

where $\alpha(n)$ is the inverse Ackermann function, a very slowly growing function. More on this see in Section 7 and 8.

Some specific configurations were investigated in previous papers (see [BGy] and [F]). The motivation of those results were geometrical.

For a matrix M (or vector as a special case) $||M||$ denotes the number of its entries equal to 1, M^T is its transpose. $[n]$ is the set of the first n positive integers, and $[a,b] =: \{a, a+1, \dots, b\}$.

2. A reduction between matrices

Let C be a configuration of 1's. We are going to define two operations on C. The first one is simply deleting an entry. The second one is attaching a new column or row to the boundary of C and placing an entry 1 in the new column or row, next to an existing one in C.

Definition 2.1. *If D can be constructed from C using one of these operations we say that D is obtained by an elementary operation from C. We use the notation $C \stackrel{e}{\rightarrow} D$. Let $\rightarrow$ be the transitive closure of $\stackrel{e}{\rightarrow}$, i.e. $C \rightarrow D$ if D can be constructed from C using a sequence of elemantary operations.*

Note that the size of the matrix can decrease by the first type of elementary operation if the deletion of the given entry creates an empty row or column.

Figure 2.2 shows several configurations and their relations.

Figure 2.2.

Theorem 2.3. *Let C, D be configurations such that $C \rightarrow D$ by t elementary steps. Then $f(n,m;D) \leq f(n,m;C) + t \cdot \max(n,m)$.*

Proof: It is sufficient to prove the case $t = 1$, we can assume that $C \stackrel{e}{\rightarrow} D$. If D is constructed by deleting an entry then the claim is obvious. So we can assume that D is constructed by adding an extra column to the end of C with an extra 1 (the other cases are very similar). Let M be a matrix of size $n \times m$ with $f(n,m;D)$ many 1's such that it doesn't have D as a subconfiguration. Let M' be the matrix that we get if we delete the last 1 in each row (assuming that there is any). Easy to realize that M' doesn't have C as a subconfiguration. So the number of remainder 1's in M' is at most $f(n,m;C)$. ∎

The natural way to apply Theorem 2.3 is that in the case of $C \rightarrow D$ an upper bound on $f(n;C)$ gives an upper bound on $f(n;D)$ and a construction for a matrix not having D as a submatrix gives a good construction for C.

Figure 2.4 contains some additional matrices with four 1's and some of their $\rightarrow$ relations.

Figure 2.4.

Let B_2 be $(1,1)$, a 1×2 configuration.

Proposition 2.5. *If $B_2 \to C$ and C has at least 2 entries in it then*

$$\min(n,m) \le f(n,m;C) \le c_C(n+m).$$

Proof: Trivial. The lower bound comes considering a matrix M with 1's only in one row or in one column.

The upper bound is immediate from Theorem 2.3. ∎

Figure 2.6.

We remark that Figure 2.2, Figure 2.4 and Figure 2.6 contain all the 37 configurations with four 1's (not distinguising two if they are the same upto rotations and reflections). The simple reduction principle yields that 22 of them have linear complexity.

Corollary 2.7.

(1) *If M has at most 3 non-zero entries then $f(n,m;M) \le 2(n+m)$.*
(2) *The 22 matrices on Figure 2.6 have linear complexity, $f(n,m;C_i) \le 3(n+m)$ for $16 \le i \le 37$.* ∎

One can extend the $\to$ relations to sets of configurations. This will be proven very useful.

Definition 2.8. *Let $C^1,\dots,C^k$ be a set of configurations. We are going to define two operations. One is simply adding a new configuration to our set. The second is substitute a C^i with D if $C^i \to D$. The transitive closure of these relations is $\to$.*

The notation is not in conflict with Definition 2.1, which is a special case of this. Note that $\{C^1,\dots,C^k\} \to \{D^1,\dots,D^l\}$ iff for every i there is a j such that $C^i \to D^j$ according to Definition 2.3.

The analog of Theorem 2.3 is the following.

Theorem 2.9. *If $\{C^1,\dots,C^k\} \to \{D^1,\dots,D^l\}$ then $f(n,m;\{D^1,\dots,D^l\}) \le f(n,m;\{C^1,\dots,C^k\}) + const(n+m)$, where the constant depends only on the two systems, and not on n and m.* ∎

A few examples:

$$\left\{\begin{pmatrix} 1 & 1 & \\ 1 & & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} 1 & & 1 \\ & 1 & 1 \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \\ 1 & 1 \end{pmatrix}\right\} \to \begin{pmatrix} & 1 & & \\ & & & 1 \\ 1 & & & \\ & & 1 & \end{pmatrix},$$

$$\left\{\begin{pmatrix} 1 & 1 & \\ 1 & & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & \\ & 1 \end{pmatrix}\right\} \to \begin{pmatrix} 1 & 1 & \\ 1 & & \\ & & 1 \end{pmatrix}.$$

3. Matrices with $n \log n$ complexity

Theorem 3.1. ([F]) $f(n, \begin{pmatrix} 1 & 1 & \\ 1 & & 1 \end{pmatrix}) < 6n \log n$. ∎

The construction in [F] shows that this upper bound is the best up to a constant factor. Below we give another, a simpler recursive construction.

Construction 3.2. *Let*

$$A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

and

$$A_{n+1} = \begin{pmatrix} E_{2^n} & A_n \\ A_n & 0_{2^n} \end{pmatrix},$$

where E_n is an $n \times n$ matrix with 1's only in the diagonal connecting the upper right to the lower left corner, and 0_n the $n \times n$ zero matrix.

Claim 3.3.

(1) *A_n is a $2^n \times 2^n$ matrix with $(n+2)2^{n-1}$ many 1's.*

(2) *A_n does not have $C_4 = \begin{pmatrix} 1 & 1 & \\ 1 & & \\ & & 1 \end{pmatrix}$ as a subconfiguration.*

Proof: (1) Easy induction.

(2) Using induction. The initial case is obvious. Let us assume that the claim is verified for A_k, when $k < n$.

Suppose on the contrary that A_n has the forbidden configuration. A_n is, by definition, divided into 4 submatrices. We distinguish different cases depending on which submatrix has the upper left corner of the forbidden configuration. If one of the A_{n-1}'s is the one, then our inductional hypothesis gives the contradiction. If $E_{2^{n-1}}$ has that entry then easy to verify that the bottom right corner of the configuration must be in $0_{2^{n-1}}$. This contradicts the fact that $0_{2^{n-1}}$ has no 1 entry at all. ∎

Corollary 3.4.

(1) $f(n; C_2)$, $f(n; \{C_2, C_2^T\})$, $f(n; C_4) = \Theta(n \log n)$.

(2) $f(n; C_i) < 10n \log n$, *for* $4 \le i \le 15$.

Proof: (1) Both the lower and upper bound comes from the following relations. $C_2 \to \{C_2, C_2^T\} \to C_4$.

(2) See Figure 2.2 and Figure 2.4. ∎

4. A CONSTRUCTION WITH $\frac{n \log n}{\log \log n}$ 1'S

In the previous section we saw an $n \log n$ upper bound on $f(C_5)$. Now we construct a matrix with $\Theta(\frac{n \log n}{\log \log n})$ 1's and not having C_5 as a subconfiguration. This section is a slightly simplified version of [BGy]. Our construction will be recursive and it defines $N(s,t)$, a matrix of size $st \times st$, where $s, t \geq 1$.

First we discuss a few properties of $N(s,t)$ what we need for the formal definition of the matrix. The st rows are divided into s blocks, each having t consecutive rows. In each block we have a column such that each of its entries are 1's and these are the first 1's in the corresponding rows. This column is the *leading column* of that block.

Let $N(s)$ be a $s \times s$ matrix without the configurations:

$$C_5 = \begin{pmatrix} 1 & & 1 \\ & 1 & \\ 1 & & \end{pmatrix}, \begin{pmatrix} 1 & & 1 \\ 1 & 1 & \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{4.1}$$

DEFINITION 4.2. *$N(1,t)$ is a $t \times t$ matrix with t 1's in the first column and 0's everywhere else.*

$N(s,1)$ is the $s \times s$ identity matrix.

The construction of $N(s,t+1)$ is the following (we assume that $N(s,t')$ and $N(s',t'')$ are already constructed for $t' \leq t$, $s' < s$ and t'' arbitrary). Take a copy of $N(s,t)$ and insert an extra row after each block. In each extra row put a 1 at the leading column of the block just above it. Add s new columns at the end of the already constructed part. At the intersection of extra rows and new columns we have an $s \times s$ space. Put a copy of $N(s)$ with maximum number of ones.

The promised propeties are maintained so our recursion is correct.

THEOREM 4.3. *$N(s,t)$ doesn't have the configurations given in (4.1).*

PROOF: An easy induction by case by case check. ∎

The previous theorem gives lower bounds on the complexity of several configurations and sets of configurations.

COROLLARY 4.4. *([BGy]) $f(n; C_5) = \Omega(\frac{n \log n}{\log \log n})$.*

PROOF: Let $f(s,t) = ||N(s,t)||$ and $f(s) = \max ||N(s)||$. We have

$$f(s, t+1) \geq f(s,t) + f(s) + s,$$

and for $s \geq ab$

$$f(s) \geq f(a,b).$$

These inequalities imply that

$$f(l^a, t) \geq (t-1)l(l^a + l^{a-1} - (l-1)^a) + l^a,$$

especially

$$f(l^{a+1}) \geq f(l^a, l) \geq l^{a+2} - l(l-1)^{a+1}.$$

Letting $a = l-1$, $n = l^l$ we obtain the desired bound. ∎

5. More matrices with linear complexity

Recall that $C_{11} =: \begin{pmatrix} 1 & & & 1 \\ & 1 & 1 & \end{pmatrix}$. In this section we prove, that the complexities of $C_{11}, \ldots, C_{15}$ are all linear, at most $9n$. As one can see from Fig. 2.4, and Theorem 2.3 the above result implied by the following theorem.

Theorem 5.1. $f(n, C_{11}) \leq 7n$.

Proof: Let $A' = (a'_{ij})$ be an $n \times m$ 0–1 matrix without C_{11}. Delete the first and the last entry in each row, and delete all entries in that row if $||(a'_{ij})_{1 \leq j \leq m}|| \leq 3$. For the obtained matrix $A = (a_{ij})$ we have

$$||A'|| \leq ||A|| + 3n. \tag{5.2}$$

A does not contain the following configurations either:

$$\begin{pmatrix} 1 & & 1 \\ & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & 1 \\ 1 & 1 & \end{pmatrix}.$$

For the i'th row (if it is non-empty) let $m(i)$ $(M(i))$ denote the minimum (maximum, resp.) index in that row, i.e. $m(i) := \min\{j : a_{ij} = 1\}$. Then $[m(i), M(i)] \subset [m(i'), M(i')]$ implies $i \leq i'$.

The element a_{ij} is called *type* α if $a_{ij} = 1$, it is not the first neither the last one in its row $(m(i) < j < M(i))$, $\alpha < i$, $j \in [m(\alpha), M(\alpha)]$, and i is minimal with respect to these constraints. By definition, there are no two entries of type α in distinct rows. But there are no two 1's of type α in the i'th row either, otherwise together with $a_{\alpha, m(\alpha)}$ and $a_{\alpha, M(\alpha)}$ they form a forbidden subconfiguration. So the number of entries in A which are

(1) first or last in their row,
(2) on the top of their column, or
(3) have a type α

is at most $4n$. We claim that all the entries of A fall one of the above 3 categories, implying $||A|| \leq 3n + m$. Then (5.2) finishes the proof of the Theorem.

Suppose that the entry $a_{ij} = 1$ is not the first or the last one in the i'th row, and that there exists a $t \leq i$ with $a_{tj} = 1$. Then $j \in [m(t), M(t)]$. Let α be the maximum index, such that $\alpha < j$, and $j \in [m(\alpha), M(\alpha)]$. Then a_{ij} has type α.

Indeed, suppose on the contrary, that some entry $a_{i'j'}$ has type α, with $\alpha < i' < i$. Then, $j \in [m(\alpha), M(\alpha)] \subset [m(i'), M(i')]$, so the existence of i' contradicts the definition of α. ∎

Let C^t be a $2 \times (t+2)$ configuration with 1's in the positions $(1,1), (1,t+2)$ and $(2,2), \dots, (2,t+1)$. $C_{11} = C^2$. Deleting from every row the *middle* $t-2$ entries, Theorem 5.1 implies

Corollary 5.3. $f(n; \begin{pmatrix} 1 & & & \cdots & & & 1 \\ & 1 & 1 & \cdots & 1 & 1 & \end{pmatrix}) = f(n; C^t) \leq (t+5)n$. ∎

Finally we mention a generalization of this idea in the direction of sequences with forbidden subsequences. The following corollary is a special case of the result in [AKV].

Corollary 5.4. *Suppose that the seqence $s = x_1, x_2, \dots, x_l$ with $x_i \in [n]$, does not have two identical consecutive members, and does not contain the subsequence abba, where $a < b$, then $l \leq 100n$.*

Proof: (*sketch*) Split s into n equal parts $s = s_1 s_2 \dots s_n$, $||s_i|| = 100$. Then there is a subset $s_i^{'} \subset s_i$ containing only distinct elements with $|s_i^{'}| \geq 9$. Put 1's into the i'th column of an $n \times n$ matrix A according to $s_i^{'}$. Finally, apply 5.1 to A to get a C_{11}, and then to get an *abba* in s. ∎

6. A covering lemma

In this section we prove a covering lemma about $0-1$ matrices. As an easy application of our lemma we get several new matrices with linear complexity.

We start with a definition. An intersection of s consecutive rows and t consecutive columns is called a *rectangle*. The horizontal size of R is t and it is denoted by $h(R)$, the vertical size of R is s and it is denoted by $v(R)$. M, itself is an example for a rectangle.

Lemma 6.1. *Let M be arbitrary $0-1$ matrix. Then there is a system of rectangles $\{R_i\}$ such that*

(1) *R_i's cover all the 1's,*
(2) *$\sum_i h(R_i) \leq 4h(M)$ and $\sum_i v(R_i) \leq 4v(M)$,*
(3) *each R_i has a 1 in the upper left or bottom right corner.*

Proof: Let us define a partial order between the positions in a given matrix. We say that $a \leq b$, if the row of a is not later than b's one and a's column is not later than b's one. a Œ b if $a \leq b$ and $a \neq b$.

There are incomparable positions. For two incomparable positions c and d we say that c Ø d if c's row is earlier than d's.

Take M and consider only the positions where we have a 1. Let m_1 Ø m_2 Ø $\dots$ Ø m_k be the set of minimal 1's for the partial order Œ. Let M_1 Ø M_2 Ø $\dots$ Ø M_l

be the set of maximal 1's for the partial order Œ. We can assume that m_1 is in the first column, m_k is in the first row, M_1 is in the last row and M_l is in the last column of M.

Let $m_{i+\frac{1}{2}}$ (for $i = 1, \ldots, k-1$) be the position in the intersection of the row of m_i and the column of m_{i+1}. Let $m_{\frac{1}{2}}$ be the lower left corner of M. Let $m_{k+\frac{1}{2}}$ be the upper right corner of M. Let $M_{j+\frac{1}{2}}$ (for $j = 1, \ldots, l-1$) be the position in the intersection of the column of M_i and the row of M_{i+1}. Let $M_{\frac{1}{2}} = m_{\frac{1}{2}}$ and $M_{l+\frac{1}{2}} = m_{k+\frac{1}{2}}$. Let $h_i = [m_i, m_{i+\frac{1}{2}}]$ be a horizontal interval of positions in the row of m_i, with endpoints at m_i and $m_{i+\frac{1}{2}}$. Let v_i be the vertical interval $[m_{i-\frac{1}{2}}, m_i]$. We define the corresponding intervals for maximal 1's. Let $V_i = [M_i, M_{i+\frac{1}{2}}]$ and $H_i = [M_{i-\frac{1}{2}}, M_i]$. It is clear that $v_1, h_1, v_2, ..., v_k, h_k$ and $H_1, V_1, H_2, V_2, ..., H_l, V_l$ defines two stair shaped curves. Let us denote them by s and S. By definition it is straightforward that there are no 1 above s and below S.

Now we are starting to construct our covering system of rectangles. This system is containing two sequences of rectangeles: $\{Q_i\}$ and $\{P_i\}$. The Q_i's are going to have an entry 1 at the bottom left corner, the P_i's are going to have a 1 at the upper left corner. We define them recursively.

Let Q_1 be a rectangle with lower right corner at M_1, with lower left corner at $m_{\frac{1}{2}}$. So its right vertical side is on the vertical half line starting at M_1, going up. The missing corner of Q_1 on this line is where it first hits s.

Q_1 might cover several h_i intervals. Let h_i the first one which is not covered by Q_1. Let P_1 be a rectangle with upper left corner at m_i. This fact gives us two half lines starting at m_i and going down and to right. They hit S at two positions. They will be two other corners of Q_1.

Next, we will explain the general step in the definition.

Let us assume that we already defined $Q_1, P_1, \ldots, Q_i, P_i$. Let V_j be the first vertical interval of S which is not covered by $Q_1 \cup \ldots \cup P_i$. Let M_j be bottom right corner of Q_{i+1}. That defines two half lines starting at M_j, one going up (let us say e_{i+1}) and one going to the left. They hit s at two positions. They give us two other corner of Q_{i+1}. This completes the definition of Q_{i+1}.

Let us assume that we already defined $Q_1, P_1, \ldots, Q_i, P_i, Q_{i+1}$. Let h_j be the first horizontal interval of s which is not covered by $Q_1 \cup \ldots \cup P_i \cup Q_{i+1}$. Let m_j be upper left corner of P_{i+1}. That defines two half lines starting at m_j, one going down and one going to the right (f_{i+1}). They hit S at two positions. They give us two other corner of P_{i+1}. This completes the definition of P_{i+1}.

The procedure stops when the already constructed rectangles cover all the V_j's (or all the h_j's).

Now we prove that the constructed system of rectangles satisfy (1)-(3).

(3) is immediate.

In order to prove (1) we need a few remarks.

It is immediate from the definition that as i is increasing the lines, e_i's are moving to the left and the lines f_i's are moving up.

The definition also implies that the upper left corner of P_i is on e_i or is left from e_i. Similarly the lower right corner of Q_{i+1} is on f_i or is below f_i. This guarantes that $Q_1 \cup \ldots \cup P_i \cup Q_{i+1}$ covers everything left from e_{i+1} in the region between s and S. Similarly $Q_1 \cup \ldots \cup P_i \cup Q_{i+1} \cup P_{i+1}$ covers everything below f_{i+1} in the region s and S. This proves (1).

For (2): From the definition the top side of Q_i (and this way the whole rectangle) is not above f_i. The lower right corner of Q_{i+1} (let us say M_j) is not above f_i, but it is the last maximal 1 with this property. This guarantees that Q_{i+2}'s lower right corner (and this way the whole rectangle) is above f_i. So the rows of Q_i and Q_{i+2} are completely disjoint. One gets the corresponding statements for the columns and for the P_i's similarly. (2) is an easy consequence of this.

This completes the proof. ∎

Corollary 6.2.

(1) $f(n,m;C_{10})$ *is linear.*

(2) $f(n,m;\begin{pmatrix} 1 & & & & \\ & & & 1 & \\ & & 1 & & \\ & 1 & & & \\ & & & & 1 \end{pmatrix})$ *is linear.*

(3) $f(n,m;\{\begin{pmatrix} 1 & 1 & \\ 1 & & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & & 1 \\ & 1 & 1 \end{pmatrix}\})$ *is linear.*

Proof: (1) Take the cover guaranted by Theorem 6.1. Count the 1's separately in different covering rectangles. We know that in the upper leftcorner or in the lower right corner there is a 1. So we can bound the number of 1's using that $f(n,m;\begin{pmatrix} 1 & & \\ & & 1 \\ & 1 & \end{pmatrix})$ and $f(n,m;\begin{pmatrix} & 1 & \\ 1 & & \\ & & 1 \end{pmatrix})$ are linear. If we add up these bounds we obtain the claim in (1).

The same proof works for (2), but there we use Theorem 5.1.

(3) follows the same way. ∎

7. Davenport-Schinzel matrices

In this and the next section we consider the complexity of $C_6 = \begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \end{pmatrix}$.

Definition 7.1. *A matrix M is called Davenport-Schinzel matrix if it does not have C_6 as a subconfiguration.*

The naming is based on the analogy between this kind of matrices and Davenport-Schinzel sequences (see [DS]).

The main result in this section is to construct a Davenport-Schinzel matrix with $\Omega(n\alpha(n))$ many 1's. Finally we discuss other configurations, missing from our matrix.

Our construction is very similar to known constructions of Davenport-Schinzel sequences (see [HS],[W]). We use the same double induction. But instead of sequences we work with matrices.

The matrices we are constructing have two parameters s and t. We refer to them as $M(s,t)$. First we describe a few properties of $M(s,t)$. The recursive definition of these matrices is assuming these properties so we need to maintain them.

(a) The size of the matrix is $tC(s,t) \times tC(s,t)$, where $C(s,t)$ is defined as follows. $C(s,t) = C(s,t-1)C(s-1, C(s,t-1))$ and $C(1,s) = 1$ and $C(s,1) = 2$, for $s > 1$.

(b) The $tC(s,t)$ many rows are divided into blocks. We will refer to them as horizontal blocks. One block contains t rows (hence we have $C(s,t)$ many blocks). Let H_i be the set of the $((i-1)t+1)^{\text{st}}$,...,$(it)^{\text{th}}$ rows, i.e. the i^{th} horizontal block.

(c) Inside H_i the appearance of the first 1 happens in the same column (considering different rows). Let us say this is the $(c_i)^{\text{th}}$ column. The 1's in these columns are called *leading 1's.*

(d) $1 = c_1 < c_2 < c_3 < \ldots < c_{C(s,t)}$. These columns divide the matrix into vertical blocks. Let V_i be the set of columns from the $(c_i)^{\text{th}}$, through $(c_{i+1}-1)^{\text{st}}$, i.e. the i^{th} vertical block.

The definition of $M(s,t)$ is going to use the matrices $S = M(s,t-1)$ and $B = M(s-1, C(s,t-1))$. (Think about S as a small matrix and about B as a big matrix.) B has $C(s-1, C(s,t-1))$ many horizontal blocks of size $C(s,t-1)$. B has $C(s-1, C(s,t-1))$ many vertical blocks too. Let v_i be the number of columns contained in the i^{th} one. S has $C(s,t-1)$ many blocks (one for each row in a block of B).

The following definition assumes properties (a)-(d). (So one must check that these properties are maintained.)

Definition 7.2. *$M(1,s)$ is an identity matrix of size $s \times s$. $M(s,1)$ is $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ (for $s > 1$).*

In order to define $M(s,t)$ take $C(s-1, C(s,t-1))$ many copies of S (one for each horizontal block of B). The construction of M will be completed in $C(s-1, C(s,t-1))$ many stages. In the i^{th} stage we add $(t-1)C(s,t-1) + C(s,t-1)$ many new rows and $(t-1)C(s,t-1) + v_i$ many new columns to the part already built. The construction starts with the empty matrix. The general (i^{th}) stage is the following.

(1) We put $(t-1)C(s,t-1)$ many new rows and new columns after the already

existing ones. In the intersection of the new rows and columns we place a copy of S.

(2) We insert an extra row *after each horizontal block of the new copy of S. In these extra rows we place one extra 1, under each leading column.*

(3) Finally we add v_i new columns (after the old ones). In the new space we place a copy of the i^{th} vertical block of B using the extra rows.

The constructed matrix $M = M(s,t)$ has properties (a)-(d).

Let us introduce a few notations. *Ordinary rows* and *ordinary columns* are the rows and columns introduced in step (1). *Extra rows* are the rows introduced in step (2). *Extra columns* are the ones introduced in step (3). The 1's introduced in step (1) are the *ordinary 1's.* The 1 entries introduced in step (2) are called the *extra 1's.* The 1's introduced in step (3) are the *new 1's.*

The previous notations give a partition of 1's into new, ordinary and extra 1's. There are similar partitions for rows and columns.

Any extra 1 is in an ordinary column and in an extra row.

The next lemma summarizes a few simple statements about the matrix $M(s,t)$.

Lemma 7.3.

(1) *If s and t are chosen appropriately and $n = sC(s,t)$ then $M(s,t)$ is an $n \times n$ matrix with $n\alpha(n)$ many 1's.*
(2) *The $(c_i)^{th}$ column contains 1's inside H_i and no other 1's.*
(3) *Inside H_i, after the leading column the 1's are decreasing, i.e. if k and l are two 1's in the same horizontal block and they are not leading 1's then k Œ l or l Œ k. (Recall that q Œ p vaguely means that p is south, east or south-east direction from q.)*
(4) *If l is a new 1 and k is a 1 such that l Ø k then k is a new 1 too. (Recall that q Ø p vaguely means p is north, east or north-east direction from q.)*
(5) *If l is an ordinary 1 and k is a 1 in l's column or in l's row then k is an ordinary 1 in the same horizontal block with the one exception when l is a leading 1 and k is the extra 1 in its column.*
(6) *If l is an extra 1 or an ordinary 1 and k is an ordinary 1 such that l Ø k then l and k is in the same horizontal block.*

Proof: For (1) we refer the reader to [HS] or [W].

The proof of (2)-(6) is easy induction following the definition of $M(s,t)$. ∎

Now we are ready to discuss the missing configurations in $M(s,t)$.

Theorem 7.4. *$M(s,t)$ does not have the following configurations: (i)* $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$,
(ii) $\begin{pmatrix} & 1 & 1 \\ 1 & & 1 \end{pmatrix}$, *(iii)* $\begin{pmatrix} 1 & \\ & 1 \\ 1 & 1 \end{pmatrix}$, *(iv)* $\begin{pmatrix} 1 & & 1 \\ 1 & 1 & \end{pmatrix}$, *(v)* $\begin{pmatrix} 1 & 1 \\ & 1 \\ 1 & \end{pmatrix}$, *(vi)* $\begin{pmatrix} & 1 & & 1 \\ 1 & & 1 & \end{pmatrix}$,

(vii) $\begin{pmatrix} 1 & & 1 \\ & 1 & \\ 1 & & \end{pmatrix}$, *(viii)* $\begin{pmatrix} & 1 & & 1 \\ & & 1 & \\ 1 & & & \end{pmatrix}$.

Proof: Each configuration in the statement has four 1's in it. Let us order these 1's. A 1 is earlier than an other if its row is earlier or if they are in the same row and it is left from the other. In the case of each configuration name the four 1's as a, b, c and d following the previously defined order.

Our proof is by induction following the definition of $M(s,t)$. The initial case is $s=1$ or $t=1$. Then the statement is clear.

The induction step is proved by contradiction. Let us assume that in $M(s,t)$ we can find four different 1's: the image of a, b, c and d, such that they obey the configuration. The individual configurations are considered separately.

(i) We distinguish cases depending on what kind of entry corresponds to c. Now on we don't distinguish a, b, c, d and their images.

Case 1: c is an extra 1. Then d is a new 1. a is in a leading column but it is not an extra 1. So a's row is an ordinary row. On the other hand $c \varnothing b$, hence (by 7.3.(4)) b is a new 1. So b's row (what is the same as a's row) is an extra row. Contradiction.

Case 2: c is a new 1. Using 7.3.(4) the whole configuration consists of new 1's. So it can be recognized inside $M(s-1, C(s,t-1))$. Contradiction with the inductional hypothesis.

Case 3: c is an ordinary 1. Using 7.3.(5) the whole configuration consists of ordinary 1's from the same horizontal block. So our configuration can be recognized in a copy of $M(s,t-1)$.

(ii) *Case 1:* c is an extra 1. Then a, b and d are new 1's. Let c' the first 1 after c in its row (that row is an extra row and c' is a new 1). Easy to check that a, b, c' and d give us a configuration C_1 or one what is the same as the original configuration. So using (i) or the inductional hypothesis we get a contradiction.

Case 2: c is a new 1. a, b, c and d are all new 1's. So our configuration is in a copy of $M(s-1, C(s,t-1))$.

Case 3: c is an ordinary 1. Using 7.3.(5) our configuration is inside a copy of $M(s,t-1)$.

(iii)-(vi) Using the same case analysis based on he bottom left 1 (what is not necessarily c).

(vii) *Case1:* d is an extra 1. Then a, b and c are ordinary 1's in the same horizontal block (using 7.3.(5) and the fact that ordinary columns and rows in the same block are consecutive ones). Then the positions of b and c are contradictory with 7.3.(3).

Case 2: d is a new 1. The same as the previous second cases.

Case 3: d is an ordinary 1. The same as the previous third cases.

(viii) *Case 1:* d is an extra 1 and c is a new 1. $c \varnothing b$ hence b is a new 1 too, in particular a's and b's row is an extra row. $d \varnothing a$ so a cannot be an extra 1.

Hence all four 1's are new except d. Move d right to the first 1 in its row. Then we obtain four new ones (hence they are in a copy of $M(s-1, C(s,t-1))$ such that their configuration is the one described in (vii) or in (viii).

Case 2: d is an extra 1 and c is an ordinary 1. Using similar arguments as before we have that all four 1's are ordinary except d and they are in the same horizontal block. Move d up by one position. We obtain four ordinary 1's (inside a copy of $M(s,t-1)$) such that their configuration is the one described in (vi) or in (viii).

Case 3: d is not an extra 1. In this case take the bottom left 1 (d) and replace it with another 1 by shifting it to the leading 1 in its row and sinking it to the bottom 1 in that column. This way we obtain the same configuration but the new d is an extra 1. That was handled in the previous cases. ∎

The previous theorem gives lower bounds on the complexity of several configurations and sets of configurations.

Corollary 7.5.

(1) $f(n;C_6) = \Omega(n\alpha(n))$,
(2) $f(n;C_8) = \Omega(n\alpha(n))$,
(3) $f(n;\{\begin{pmatrix} & 1 & 1 \\ 1 & & 1 \end{pmatrix}, \begin{pmatrix} 1 & \\ & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & & 1 \\ 1 & 1 & \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ & 1 \\ 1 & \end{pmatrix}\}) = \Omega(n\alpha(n))$. ∎

8. Upper bound on Davenport-Schinzel matrices

In this section we prove that

Theorem 8.1. $f(n; \begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \end{pmatrix}) \leq O(n\alpha(n))$.

Proof: Let $A' = (a'_{ij})$ be an $n \times m$ 0–1 matrix not having a subconfiguration of $C_6 = \begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \end{pmatrix}$. Delete the first and the last 1 in each row, and keep only the columns with at least 2 entries. The obtained matrix is denoted by $A = (a_{ij})$, and obviously, for the number of entries we have $||A'|| \leq ||A|| + 2n + m$. Form a sequence s_j from the j'th column $(a_{ij})_{1\leq i\leq n}$ of length $||(a_{ij})|| =: l(j)$ in the following way

$$s_j = (s^j_1,\ s^j_2, \ldots, s^j_{l(j)}),$$

where $s^j_1 < s^j_2 < \cdots < s^j_{l(j)}$ and $a_{s_i j} = 1$ for $1 \leq i \leq l(j)$. Form one sequence $s' =: s_1 s_2 \ldots s_m$ in this order. Delete from s' the element $s^j_{l(j)}$ if it equals to s^{j+1}_1. In the obtained sequence, s, there are no equal consequtive elements. We claim that s does not contain a subsequence $ababa$, i.e. it is a $DS_5(n)$ sequence.

Suppose on the contrary. Then there exists a subsequence $abab$ of s with $a < b$. So there are $j_1 \leq \cdots \leq j_4$ such that $a \in s_{j_1}$, $b \in s_{j_2}$, $a \in s_{j_3}$, $b \in s_{j_4}$. Here

$j_2 < j_3$, otherwise the first b in $abab$ could not preceed the second a in s. Consider the submatrix defined by the rows a and b and the columns $\{j_1, \dots, j_4\}$. There are four possibilities.

$j_1 < j_2 < j_3 < j_4$

$$\begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \end{pmatrix}$$

$j_1 = j_2 < j_3 < j_4$

$$\begin{pmatrix} 1 & 1 & \\ 1 & & 1 \end{pmatrix}$$

$j_1 = j_2$ and $j_3 = j_4$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$j_1 < j_2 < j_3 = j_4$

$$\begin{pmatrix} 1 & & 1 \\ & 1 & 1 \end{pmatrix}.$$

In each cases A will contain a C_6, a contradiction.

So $||A|| \le 2n + 2m + ds_5(n)$. ∎

COROLLARY 8.2. $f(n; C_7), f(n; C_9) = O(n\alpha(n))$. ∎

In the very same way we can obtain the following theorem. Let C^{2k} be a partial $2 \times 2k$ matrix with $c_{1,2i-1} = 1$, $c_{2,2i} = 1$ for $1 \le i \le k$.

THEOREM 8.3. $f(n; C^{2k}) \le O(ds_{4k-3}(n))$. ∎

It is not difficult to give a lower bound for $f(n; C^{2k})$ which is probably closer to f as the upper bound.

THEOREM 8.4. $f(n; C^{2k}) \ge \Omega\left(\dfrac{ds_{2k+1}(n)}{\alpha(\alpha(n))^{O(\alpha(\alpha(n))^{2k-4})}}\right)$.

REMARK: Here the right hand side is superlinear. For the best bound on $ds_{2k+1}(n)$ see [ASS].

PROOF: Let s be a Davenport-Schinzel sequence $s \in DS_{2k+1}(n)$ of length $ds_{2k+1}(n)$ such that the element i appears earlier than j for $i < j$. It is well-known [S], that

$$ds_{2k+1}(n) = O(n\alpha(n)^{O(\alpha(n)^{2k-4})}).$$

Split s into n almost equal parts $s = s_1 \dots s_n$. Let $s_i^{'}$ be a set of distinct values of s_i, $||s_i^{'}|| = x$. We have that

$$y =: \frac{ds_{2k+1}(n)}{n} - 1 \le ||s_i|| \le xO(\alpha(x)^{O(\alpha(x)^{2k-4})}) \le xO(\alpha(y)^{O(\alpha(y)^{2k-4})}).$$

Here $y = O(\alpha(n)^{O(\alpha(n)^{2k-4})})$, so $\alpha(y) = \alpha(\alpha(n)) + O(1)$.

Finally, forming the i^{th} column of an $n \times n$ matrix A from $s_i^{'}$ we obtain the desired configuration without C^{2k}. ∎

9. Conclusions and open problems

The next table summarizes our results.

Configurations	Lower bound	Upper bound
$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$\Theta(n^{\frac{3}{2}})$ see Theorem 1.1.	
$\begin{pmatrix} 1 & 1 & \\ 1 & & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 & \\ 1 & & \\ & & 1 \end{pmatrix}$	$\Theta(n \log n)$ see Theorem 3.1. and Corollary 3.4.	
$\begin{pmatrix} 1 & & \\ & 1 & \\ 1 & & 1 \end{pmatrix}$	$\Omega(\frac{n \log n}{\log \log n})$ see Corollary 4.4.	$O(n \log n)$ see Corollary 3.4.
$\begin{pmatrix} 1 & & 1 \\ 1 & & \\ & 1 & \end{pmatrix}$	$\Omega(n)$	$O(n \log n)$ see Corollary 3.4.
$\begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & & & \\ & & 1 & \\ & 1 & & 1 \end{pmatrix}$	$\Theta(n\alpha(n))$ see Corollary 7.5. and Theorem 8.1.	
$\begin{pmatrix} & 1 & & \\ & & & 1 \\ 1 & & 1 & \end{pmatrix}$, $\begin{pmatrix} & 1 & & \\ & & & 1 \\ 1 & & & \\ & & 1 & \end{pmatrix}$	$\Omega(n)$	$O(n\alpha(n))$ see Corollary 8.2.
All the other 28 matrices with 4 entries	$\Theta(n)$	

Finally we mention several open problems. The first few ones are suggested by the previous table. Even in the case of configurations with four 1's there are several unknown complexities.

Is it true that the complexity of all permutation configurations are linear?

What is the characterization of configurations with linear complexity? In extremal graph theory the forbidden subgraphs with linear threshold are exactly the trees.

Is it true, that if G is the (bipartite) graph corresponding the configuration C then

$$f(n;C) < O(T(n;G)\log n)? \tag{9.1}$$

Does (9.1) hold at least for trees?

There are several combinatorial structures with an underlying order where the similar extremal question is interesting. An example is a set of intervals on a given line. How many intervals (over n endpoints) guarantee the existence of a given interval configuration? Similar question can be asked about diagonals in a cycle. Davenport and Schinzel's original question can be extended to arbitrary forbidden subsequence. As far we know there is no organized account of these questions.

10. Acknowledgments

The authors are grateful to Márió Szegedy for his valuable remarks and suggestions.

References

[A] W. Ackermann, Zum Hilbertschen Aufbau der reelen Zahlen, *Math. Ann.*, **99** (1928), 118-133.

[AKV] R. Adamec, M. Klazar and P. Valtr, Forbidden words, *Preprint*, Department of Mathematics, Karlovy University, Prague, Czechoslovakia.

[ASS] P. Agarwal, M. Sharir and P. Shor, Sharp upper and lower bounds on the length of general Davenport-Schinzel sequences, *Preprint.*

[Bo] B. Bollobás, *Extremal Graph Theory*, Academic Press, London - New York, 1978.

[B] W.G. Brown, On graphs that do not contain a Thomsen graph, *Canad. Math. Bull.*, **9** (1966), 281-285.

[BGy] D. Bienstock and E. Győri, An extremal problem on sparse $0-1$ matrices, to appear in *SIAM J. Disc. Math.*

[DS] H. Davenport and A. Schinzel, A combinatorial problem connected with differential equations, I and II, *Amer. J. Math.*, **87** (1965), 684-694 and *Acta Arithmetica*, **17** (1971), 363-372.

[ERS] P. Erdős, A Rényi and V.T. Sós, On a problem of graph theory, *Studia Sci. Math. Hungar.*, **1** (1966), 215-235.

[ESi] P. Erdős and M. Simonovits, A limit theorm in graph theory, *Studia Sci. Math. Hungar.*, **1** (1966), 51-57.

[ESt] P. Erdős and A.H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.*, **52** (1946), 1087-1091.

[F] Z. Füredi, The maximum number of unit distances in a convex n-gon, to appear in *J. Combinatorial Th., A.*

[HS] S. Hart and M. Sharir, Nonlinearity of Davenport-Schinzel sequences and of generalized path compression schemes, *Combinatorica*, **6** (1986), 151-177.

[K] P. Komjáth, A simplified construction of nonlinear Davenport-Schinzel sequences, *J. of Comb. Theory*, **A 49** (1988), 262-267.

[KST] T. Kővári, V.T. Sós and P. Turán, On a problem of Zarankiewicz, *Colloq. Math.*, **3** (1954), 50-57.

[M] J. Mitchell, Shortest rectilinear paths among obstacles, *SORIE Technical report No. 739*, Cornell University, 1987.

[S] M. Sharir, Almost linear upper bounds on the length of generalized Davenport-Schinzel sequences, *Combinatorica*, **7** (1987), 131–143.

[Sz] E. Szemerédi, On a problem by Davenport and Schinzel, *Acta Arithmetica*, **15** (1974), 213–224.

[W] A. Wiernik, Planar realizations of Nonlinear Davenport-Schinzel sequences by segments, in *Procedings, 27th IEEE Found. of Comput. Sci.*, (1986), 97-106; to appear in Discret and Comput. Geom.

$$C_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\downarrow$$

$$C_2 = \begin{pmatrix} 1 & 1 & \\ 1 & & 1 \end{pmatrix}$$

$$\swarrow \qquad\qquad \downarrow \qquad\qquad \searrow$$

$$C_3 = \begin{pmatrix} & 1 & \\ 1 & & \\ 1 & & 1 \end{pmatrix} C_4 = \begin{pmatrix} 1 & 1 & \\ 1 & & \\ & & 1 \end{pmatrix} C_5 = \begin{pmatrix} 1 & & \\ & 1 & \\ 1 & & 1 \end{pmatrix}$$

$$C_6 = \begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \end{pmatrix}$$

$$\swarrow \searrow$$

$$C_7 = \begin{pmatrix} & & 1 & \\ 1 & & & \\ & 1 & & 1 \end{pmatrix} C_8 = \begin{pmatrix} 1 & & & \\ & & 1 & \\ & 1 & & 1 \end{pmatrix}$$

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

$$C_9 = \begin{pmatrix} & & 1 & \\ 1 & & & \\ & & & 1 \\ & 1 & & \end{pmatrix} C_{10} = \begin{pmatrix} 1 & & & \\ & & 1 & \\ & 1 & & \\ & & & 1 \end{pmatrix}$$

Figure 2.2.

$$
\begin{tikzcd}[column sep=small]
& C_{11} = \begin{pmatrix} 1 & & & 1 \\ & 1 & 1 & \end{pmatrix} \arrow[dl] \arrow[dr] & \\
C_{12} = \begin{pmatrix} 1 & & & 1 \\ & 1 & & \\ & & 1 & \end{pmatrix} \arrow[d] & & C_{13} = \begin{pmatrix} 1 & & & \\ & & & 1 \\ & 1 & 1 & \end{pmatrix} \arrow[d] \\
C_{14} = \begin{pmatrix} & & & 1 \\ 1 & & & \\ & 1 & & \\ & & 1 & \end{pmatrix} & & C_{15} = \begin{pmatrix} 1 & & & \\ & & & 1 \\ & 1 & & \\ & & 1 & \end{pmatrix}
\end{tikzcd}
$$

Figure 2.4.

$$C_{16}=\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&1\end{pmatrix},C_{17}=\begin{pmatrix}1&&&\\&1&&\\&&&1\\&&1&\end{pmatrix},C_{18}=\begin{pmatrix}&1&&\\1&&&\\&&&1\\&&1&\end{pmatrix},$$

$$C_{19}=\begin{pmatrix}&1&&\\1&&&1\\&&1&\end{pmatrix},C_{20}=\begin{pmatrix}1&1&&\\&&1&\\&&&1\end{pmatrix},C_{21}=\begin{pmatrix}1&1&&\\&&&1\\&&1&\end{pmatrix},$$

$$C_{22}=\begin{pmatrix}1&&&\\&1&1&\\&&&1\end{pmatrix},C_{23}=\begin{pmatrix}1&&&\\&&1&1\\&1&&\end{pmatrix},C_{24}=\begin{pmatrix}1&&&\\&1&&1\\&&1&\end{pmatrix},$$

$$C_{25}=\begin{pmatrix}1&1&\\&1&\\&&1\end{pmatrix},C_{26}=\begin{pmatrix}1&1&\\&&1\\&1&\end{pmatrix},C_{27}=\begin{pmatrix}1&1&\\&&1\\&&1\end{pmatrix},$$

$$C_{28}=\begin{pmatrix}1&&1\\&1&\\&1&\end{pmatrix},C_{29}=\begin{pmatrix}1&&\\&1&1\\&1&\end{pmatrix},C_{30}=\begin{pmatrix}&1&\\1&&1\\&1&\end{pmatrix},$$

$$C_{31}=\begin{pmatrix}1&1&1&\\&&&1\end{pmatrix},C_{32}=\begin{pmatrix}1&1&&\\&&1&1\end{pmatrix},C_{33}=\begin{pmatrix}1&&1&1\\&1&&\end{pmatrix},$$

$$C_{34}=\begin{pmatrix}1&1&1\\1&&\end{pmatrix},C_{35}=\begin{pmatrix}1&1&1\\&1&\end{pmatrix},C_{36}=\begin{pmatrix}1&1&\\&1&1\end{pmatrix},$$

$$C_{37}=\begin{pmatrix}1&1&1&1\end{pmatrix}.$$

Figure 2.6.