Charles University in Prague
Faculty of Mathematics and Physics

DOCTORAL THESIS

Josef Cibulka

Extremal combinatorics of matrices, sequences and sets of permutations

Department of Applied Mathematics

Supervisor of the doctoral thesis: Doc. RNDr. Pavel Valtr, Dr.
Study programme: Computer Science
Specialization: Discrete Models and Algorithms

Prague 2013

Acknowledgements

First of all, I would like to thank my advisor, Pavel Valtr for his support and advice during my studies and for negotiating several stays abroad. He brought to my attention many interesting problems, some of which evolved into results of this thesis. He also, together with Jan Kratochvíl, led a research seminar for undergraduate students where the very first of these results and also my first encounters with research originate. Many thanks to Jan Kynčl for fruitful discussions and a lot of advice. The joint results with Jan form, in my opinion, the most interesting part of this thesis.

Part of the work was done at Centre Interfacultaire Bernoulli at l'École Polytechnique Fédérale de Lausanne and at the DIMACS institute of the Rutgers University whose support is highly acknowledged.

I would also like to thank all the colleagues at the Department of Applied Mathematics and the Computer Science Institute of Charles University for a nice atmosphere. In particular, thanks to the "computer wizards" for taking care of the computers; the data obtained using the computers in many cases showed the right direction of search.

Many thanks to all my coauthors and collaborators. Thanks also to Xavier Goaoc, Vítek Jelínek and Martin Klazar for discussions about problems covered in this thesis and to Ida Kantor for introducing me to a problem that lead to one of the results.

Last but not least I would like to thank my family for their support.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In Prague, 15th January 2013 Josef Cibulka

Název práce: Extremální kombinatorika matic, posloupností a množin permutací

Autor: Josef Cibulka

Katedra: Katedra aplikované matematiky

Vedoucí disertační práce: Doc. RNDr. Pavel Valtr, Dr., Katedra aplikované matematiky

Abstrakt: V této práci se zabýváme oblastmi extremální teorie $\{0,1\}$-matic, posloupností a množin permutací, které mají četná využití v oblasti kombinatorické a výpočetní geometrie. *VC-dimenze* množiny n-prvkových permutací $\mathcal{P}$ je největší celé číslo k takové, že množina zúžení permutací z $\mathcal{P}$ na některou k-tici pozic je množina všech k-prvkových permutací. Projdeme všemi třemi zmíněnými oblastmi extremální kombinatoriky, abychom dokázali horní a dolní meze, rostoucí kvaziexponenciálně v n, na maximální možnou velikost množiny n-permutací s VC-dimenzí shora omezenou konstantou. Tento výsledek využívá ve svém článku Jan Kynčl k výraznému snížení horního odhadu na počet tříd slabého izomorfismu úplného topologického grafu na n vrcholech. Dále pro některé, zejména permutační, matice M dokážeme nové meze na počet jedniček v M-prosté $\{0,1\}$-matici velikosti $n \times n$. Například pro každé k zkonstruujeme matici s $k^2n/2$ jedničkami prostou jedné konkrétní permutační matice velikosti $k \times k$. Také dokážeme téměř těsné meze na maximální počet jedniček v matici prosté pevně zvolené vrstvené permutační matice.

Klíčová slova: extremální teorie, zakázaná podstruktura, množina permutací, $\{0,1\}$-matice

Title: Extremal combinatorics of matrices, sequences and sets of permutations

Author: Josef Cibulka

Department: Department of Applied Mathematics

Supervisor: Doc. RNDr. Pavel Valtr, Dr., Department of Applied Mathematics

Abstract: This thesis studies questions from the areas of the extremal theory of $\{0,1\}$-matrices, sequences and sets of permutations, which found many applications in combinatorial and computational geometry. The *VC-dimension* of a set $\mathcal{P}$ of n-element permutations is the largest integer k such that the set of restrictions of the permutations in $\mathcal{P}$ on some k-tuple of positions is the set of all $k!$ permutation patterns. We show lower and upper bounds quasiexponential in n on the maximum size of a set of n-element permutations with VC-dimension bounded by a constant. This is used in a paper of Jan Kynčl to considerably improve the upper bound on the number of weak isomorphism classes of complete topological graphs on n vertices. For some, mostly permutation, matrices M, we give new bounds on the number of 1-entries an $n \times n$ M-avoiding matrix can have. For example, for every even k, we give a construction of a matrix with $k^2n/2$ 1-entries that avoids one specific k-permutation matrix. We also give almost tight bounds on the maximum number of 1-entries in matrices avoiding a fixed layered permutation matrix.

Keywords: extremal theory, forbidden substructure, set of permutations, $\{0,1\}$-matrix

Contents

Introduction

Overview and motivation

This thesis considers problems from the area of extremal combinatorics. We mainly focus on the extremal theory of sequences, $\{0, 1\}$-matrices and sets of permutations.

In this introduction, for each of these areas, we start with the necessary definitions, mention the most important results and then describe some applications in discrete geometry. The last part of the introduction is devoted to the description of the main results of this thesis.

These areas of extremal combinatorics turn out to be closely related to each other and many results of this thesis provide a modest contribution to the connections between them. Another area connected to the problems discussed below is the Turán theory of graphs. Results on Dirac problems on graphs and from the extremal set theory are also applicable in the extremal theory of sets of permutations; as shown by Cheong, Goaoc and Nicaud [CGN13].

The thesis is based on the following four papers.

(1) Josef Cibulka. On constants in the Füredi–Hajnal and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory, Series A*, 116(2):290–302, 2009.

(2) Josef Cibulka. Extremal functions of linear forbidden matrices. In preparation.

(3) Josef Cibulka and Jan Kynčl. Tight bounds on the maximum size of a set of permutations with bounded VC-dimension. *J. Comb. Theory, Ser. A*, 119(7):1461–1478, 2012.

(4) Josef Cibulka. Maximum size of reverse–free sets of permutations. To appear in *SIAM Journal on Discrete Mathematics*. Preprint available at: `http://arxiv.org/abs/1208.2847`.

Chapters 3 and 4 are slightly modified versions of papers 3 and 4, respectively. Sections 2.2, 2.3, 1.4 and 2.5 are based on the paper 1. The rest of Chapters 1 and 2 will be submitted as paper 2.

Generalized Davenport–Schinzel sequences

A real function in variable n is *quasilinear* if it is of the form $n\beta(n)$, where $\beta(n)$ is defined in terms of the inverse Ackermann function, which grows extremely slowly. Similarly, the functions of the form $2^{n\beta(n)}$ are *quasiexponential*. Since these notions are vague, the interested reader is advised to have a look at examples of quasilinear functions in Theorem 3.3 and of quasiexponential functions in Theorem 3.2. The definition of the inverse Ackermann function can be found in Section 3.2.2.

Let $[a, b]$ be the set of integers $\{a, a+1, \ldots, b\}$ and let $[n]$ be the set of integers from 1 to n, that is, $[n] = [1, n]$. A *sequence* (or a *word*[1]) w of length k over the alphabet Γ is a sequence $(w_1, \ldots, w_k)$ of elements from Γ. The set of all words of length k over $[n]$ is $[n]^k$. An *n-permutation* is a word from $[n]^n$ where every element of $[n]$ appears exactly once. In the case of sequences over small alphabets, the parentheses and commas are often left out. That is, 1324 represents the permutation $(1, 3, 2, 4)$.

Let S and T be sequences. We say that S *contains* a pattern T if S contains a subsequence T' isomorphic to T, that is, T can be obtained from S by a removal of some of its letters followed by a one-to-one renaming of the symbols. A sequence S over an alphabet Γ is a *Davenport–Schinzel sequence of order s* (DS(s)-sequence for short) if no symbol appears on two consecutive positions and S does not contain the alternating pattern $abab\ldots$ of length $s+2$. For example, the longest alternating pattern in 12352342 has length 5 and is formed by symbols 2 and 3. Both the lower and upper bounds on the maximum length of a Davenport–Schinzel sequence of order $s \geq 3$ are known to be quasilinear.

These sequences were introduced by Davenport and Schinzel [DS65] in the study of complexity of lower envelopes of functions. Let $f_1, \ldots, f_n$ be real functions of one variable. Their *lower envelope* is the real function whose value at every $x \in \mathbb{R}$ is the minimum of all $f_i(x)$. When the functions are continuous, then their lower envelope is continuous and the real line is divided into intervals within each of which one of the functions attains the minimum. If, in addition, each two of the functions are equal in at most s points, then the sequence formed by writing the indices of the minimal functions is a DS-sequence of order s. The study of these sequences later found numerous other applications in computational and combinatorial geometry. See, for example, the book of Sharir and Agarwal [SA95].

If a sequence avoids a repetition of the same letter within every substring of r consecutive letters, then the sequence is *r-sparse* (such sequences are also often called *r-regular*). A *generalized DS-sequence* is a sequence that is r-sparse and where an arbitrary pattern T having r distinct letters is forbidden.

Valtr [Val97] proved an $O(n \log(n))$ upper bound on the number of edges of a graph whose edges can be drawn as x-monotone curves so that every pair of edges crosses at most once and no k edges are pairwise crossing. The result uses the linear upper bound on the length of the generalized DS-sequence avoiding the "N-shaped" pattern $N_k = 12\ldots k(k-1)\ldots 212\ldots k$ [KV94]. Fox, Pach and Suk [FPS13] generalized this result by allowing pairs of edges to cross more than once. They also use the result on generalized DS-sequences.

Pettie [Pet11c] used a result on generalized DS-sequences to improve the upper bound on the complexity of the union of δ-fat triangles.

[1] The distinction between the terms *sequence* and *word* is mainly historical. We follow the tradition and use the term Davenport–Schinzel sequence, and the term word for an element of a code in Chapter 4. To avoid a total confusion, the term *function from* $[k]$ *to* $[n]$ used in the paper [CK12] was changed to the term *word* in Chapter 3.

Extremal Problems on Matrices

An $m \times n$ $\{0,1\}$*-matrix* $A = (a_{i,j})$ is a matrix from $\{0,1\}^{m \times n}$, that is, every entry is either 0 or 1. All matrices in this thesis are $\{0,1\}$-matrices; even there, where it is not explicitly mentioned. We often use the notation in which 1's are represented by dots and 0's by empty space. For example, $\left(\begin{smallmatrix}\bullet & & & \bullet\\ & \bullet & \bullet & \end{smallmatrix}\right)$ is $\left(\begin{smallmatrix}1&0&0&1\\0&1&1&0\end{smallmatrix}\right)$. An *empty row (column)* is a row (column) with no 1-entry. A matrix is *nonempty* if it has at least one 1-entry.

A $\{0,1\}$-matrix A is said to be a *permutation matrix*, if each row and each column contains exactly one 1-entry. Each such matrix corresponds to some permutation. An *n-permutation matrix* is an $n \times n$ permutation matrix, that is, it corresponds to an n-permutation.

We say that B is a *submatrix* of A if it can be obtained from A by removing some of its rows and columns. A $\{0,1\}$-matrix A *contains* a $k \times \ell$ $\{0,1\}$-matrix M if A has a $k \times \ell$ submatrix B that has 1's on all the positions where M does. For example, every matrix with at least k rows and k columns contains the empty $k \times k$ matrix. A matrix A *avoids* M if it does not contain M. Roughly speaking, forbidding a matrix corresponds to forbidding a subgraph in a bipartite graph with an additional ordering on the vertices.

The *extremal function of a forbidden matrix* M is the maximum number $\mathrm{ex}_M(m,n)$ of 1-entries in an $m \times n$ $\{0,1\}$-matrix avoiding M. Let $\mathrm{ex}_M(n) := \mathrm{ex}_M(n,n)$. Apparently, the extremal function is bounded from above by mn. When the forbidden matrix is nonempty and has at least 2 rows and 2 columns, then the extremal function is at least linear in $m+n$.

The asymptotics of the extremal function are known only for several small matrices and some specific classes of matrices. For example, forbidding the $k \times \ell$ matrix with all entries being 1 corresponds to the Turán-type problem of forbidding the complete bipartite subgraph in a bipartite graph. Assuming $k \leq \ell$, the extremal function is thus in $O(n^{2-1/k})$ and, when $k \leq 3$, in $\Omega(n^{2-1/k})$, see for example [Mat02, Chapter 4.5]. A *word matrix* is a matrix with exactly one 1-entry in every column. As a consequence of the results on generalized DS-sequences, the extremal function of every word matrix is at most quasilinear.

We call a forbidden matrix M *linear* if its extremal function is linear in $n+m$. A breakthrough in the study of linear extremal matrices occurred when Marcus and Tardos [MT04] proved that all permutation matrices are linear.

Extremal functions of word matrices often turn out to be tightly connected to bounds on the length of appropriate generalized DS-sequences. The connection between matrices with forbidden word matrices and DS-sequences was first noticed by Füredi and Hajnal [FH92]. We make use of this connection in Chapter 3, where we use quasilinear upper and lower bounds on the length of sequences satisfying conditions similar to those of generalized DS-sequences.

Pettie [Pet11c, Pet11b] went even further. He proves results both about matrices and sequences at the same time, by switching between the two areas to simplify the proofs. One of his results is an improvement on the upper bound of the generalized DS-sequence avoiding the pattern N_k [Pet11c]. This improves the dependance on k of the above mentioned bound on the number of edges of a graph with x-monotone edges and no k pairwise crossing edges. The bound is exponential in k^2. Pettie uses as a base case an upper bound on the extremal function of $\left(\begin{smallmatrix}\bullet & & & \bullet\\ & \bullet & \bullet & \end{smallmatrix}\right)$.

Results about forbidden matrices found several applications in combinatorial and computational geometry. Here we describe two of them, some more can be found in the papers of Bienstock and Győri [BG91], Pach and Tardos [PT06] and Pach and Sharir [PS91]. Füredi showed that given an n-tuple of points in convex position, the number of pairs of points at unit distance is at most $O(n \log n))$. This follows from the $\Theta(n \log(n))$ bound on the extremal function of $\begin{pmatrix} \bullet & & \bullet \\ \bullet & \bullet & \end{pmatrix}$. Efrat and Sharir [ES96] used the same forbidden matrix to give an $O(n^{1+\varepsilon})$ time algorithm for the planar segment cover problem. Given a set of n points in the plane and a real number ℓ, the *planar segment center cover problem* is to draw a straight-line segment of length ℓ that minimizes the maximum distance from a point of the given set.

Enumerative and Extremal Problems on Sets of Permutations

One permutation avoids another permutation if and only if the same holds for the corresponding permutation matrices. For example, 31254 contains 123 since it contains an increasing subsequence of length 3. Chapter 2 considers the enumeration of permutations avoiding a given permutation. The study of such questions started when Knuth [Knu68] showed that permutations sortable through a stack are exactly those that avoid 231 and that their numbers equal the Catalan numbers.

It is known [MT04, Kla00] that the number of n-permutations avoiding a given permutation π is at most s_π^n for some constant s_π. Equivalently, the number of n-permutation matrices avoiding a given permutation matrix P is at most s_P^n for some constant s_P. This was proved from the linearity of forbidden permutation matrices.

Motivated by the so-called acyclic linear orders problem, Raz [Raz00] defined the VC-dimension of a set $\mathcal{P}$ of permutations: The *restriction of a permutation π to a k-tuple of positions* is the k-permutation corresponding to the ordering of the k elements on the selected positions in π. A k-tuple of positions is *shattered by a set $\mathcal{P}$ of permutations* if each k-permutation appears as a restriction of some $\pi \in \mathcal{P}$ to the k-tuple of positions. The *VC-dimension of $\mathcal{P}$* is the size of the largest set of positions shattered by $\mathcal{P}$. Raz showed that a set of n-permutations with VC-dimension 2 has size at most C^n for some absolute constant C.

A *topological graph* is a graph drawn in the plane so that each pair of edges has at most one point in common and no three edges have a common crossing point. Two topological graphs are *weakly isomorphic* if they correspond to the same abstract graph and have the same set of pairs of crossing edges. In Chapter 3, for every fixed $k \geq 3$, we show a quasiexponential upper bound on the size of a set of n-permutations with VC-dimension k. A *topological graph* is a graph drawn in the plane so that vertices are represented by points and edges by continuous curves. Kynčl [Kyn12] uses this upper bound to significantly improve the upper bound on the number of weak isomorphism classes of simple complete topological graphs. Roughly speaking, Kynčl shows that the set of orders in which a newly added vertex can see the vertices of a complete topological graphs has VC-dimension at most 4^{30^4}.

Summary of the main results

Chapter 1 studies extremal functions of linear forbidden matrices. In particular, we are interested in the dependance of the extremal function on the size of the forbidden matrix for some classes of forbidden matrices.

In Section 1.2, we find the exact extremal function for all but one linear matrix with four 1-entries and no empty row and column. In particular, we improve the upper bound on the extremal function of $\begin{pmatrix} \bullet & & & \bullet \\ & \bullet & \bullet & \end{pmatrix}$. This slightly decreases the base of the exponent in the earlier mentioned upper bound on the length of a generalized DS-sequence avoiding the pattern N_k.

A *layered permutation* is a concatenation of decreasing sequences, called layers, such that each entry of one layer is smaller than entries of the following layers. In Section 1.3, we find tight bounds on the extremal function $\mathrm{ex}_L(n)$ for permutation matrices L corresponding to the layered permutations. The upper and lower bounds on $\mathrm{ex}_L(n)$ differ only in smaller order terms.

In Section 1.4, we find, for every even k, a k-permutation matrix $\mathrm{Cross}(k)$ whose extremal function is at least $nk^2/2$. This is the largest known lower bound on the extremal function of a k-permutation matrix.

Chapter 2 considers the enumeration of permutations avoiding a given permutation. We show that the values of s_P and $c_P = \lim_{n\to\infty} \mathrm{ex}_P(n)/n$ are closely related. We show that for every k-permutation matrix P, $s_P \leq 2.88c_P^2$ (Section 2.2), $c_P \leq \alpha s_P^{4.5}$ for some absolute constant α (Section 2.3) and $c_P \leq 8s_{P'}$, where P' is a permutation matrix of size at most $3k/2 \times 3k/2$ (Section 2.4). Section 2.5 introduces generalizations to higher-dimensional permutation matrices.

In Chapter 3, for every fixed $k \geq 3$, we show a quasiexponential upper bound on the size of a set of n-permutations with VC-dimension k. We also show relatively tight quasiexponential lower bounds. These bounds use the quasilinearity of the extremal function of matrices avoiding a word matrix and bounds on matrices satisfying a related condition.

Two words have a *reverse* if they have the same pair of distinct letters on the same pair of positions, but in reversed order. Füredi, Kantor, Monti and Sinaimeri [FKMS10] defined a *reverse-free code* to be a set of words of length n no two of which have a reverse. In Chapter 4, we show lower and upper bounds on the maximum size of reverse-free codes whose elements are permutations. Both bounds are of the form $n^{n/2+O(n/\log n)}$ and are generalized to codes made of words such that no letter occurs more than once within a word. These bounds use the $\Theta(n^{3/2})$ bound on the extremal function of $\begin{pmatrix} \bullet & \bullet \\ \bullet & \bullet \end{pmatrix}$.

1. Matrices avoiding a given matrix

1.1 Introduction

A $(0,1)$-matrix $A = (a_{i,j})$ of size $n \times n$ is said to be an *n-permutation matrix*, if each row and each column contains exactly one 1-entry. Each such matrix can be matched to some n-permutation $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ in such a way, that $A_{i,j} = 1$ exactly if $\pi_j = i$. We let P_π denote the permutation matrix corresponding to π.

A $\{0,1\}$-matrix A *contains* a $k \times k$ $\{0,1\}$-matrix $M = (M_{i,j})$ if A has a $k \times k$ submatrix $B = (B_{i,j})$ such that for all $i, j \in [k]$, $M_{i,j} = 1$ implies $B_{i,j} = 1$. A matrix A *avoids* M if it does not contain M. Note that a permutation matrix A contains another permutation matrix P if and only if P is a submatrix of A.

Given a $\{0,1\}$-matrix M, let $\mathrm{ex}_M(m,n)$ be the maximum number of 1-entries in an $m \times n$ $\{0,1\}$-matrix avoiding M. Let $\mathrm{ex}_M(n) := \mathrm{ex}_M(n,n)$. The function ex_M is the *extremal function of the forbidden matrix M*. Given a set $\mathcal{M}$ of $\{0,1\}$-matrices, let $\mathrm{ex}_{\mathcal{M}}(m,n)$ be the maximum number of 1-entries in an $m \times n$ $\{0,1\}$-matrix avoiding all matrices of $\mathcal{M}$.

The asymptotics of the extremal function are known only for specific classes of forbidden matrices.

We call a forbidden matrix M *linear* if its extremal function is linear in $n+m$. In this chapter, we study the extremal function of linear forbidden matrices.

We define the *Füredi–Hajnal limit* of M as follows:

$$c_M := \lim_{n\to\infty} \frac{\mathrm{ex}_M(n)}{n}.$$

A small modification of the proof of Theorem 1 from [Arr99] shows that whenever M is a linear forbidden matrix, then c_M exists, is finite and satisfies

$$\forall n \in \mathbb{N} : \mathrm{ex}_M(n) \le c_M n.$$

In 1992 Füredi and Hajnal [FH92] conjectured that for any fixed permutation matrix P, $\mathrm{ex}_P(n) = O(n)$. The conjecture was solved by Marcus and Tardos [MT04].

Theorem 1.1 (Füredi–Hajnal conjecture; Marcus and Tardos, 2004)**.** *For every k-permutation matrix P, there is a constant $c_P \le 2k^4\binom{k^2}{k}$ such that*

$$\mathrm{ex}_P(n) \le c_P n.$$

Let c_k be the maximum of the Füredi–Hajnal limit over all k-permutation matrices. The upper bound of Marcus and Tardos remains the best known upper bound on c_k, that is

$$c_k \le 2k^4\binom{k^2}{k}.$$

It is easy to find a general lower bound that turns out to be attained by many linear matrices. The following observation was made by several authors; see for example [Cib09, Ful09].

Claim 1.2. *Let M be a $k \times \ell$ matrix with at least one 1-entry. For any $m \geq k-1$ and $n \geq \ell - 1$:*

$$\mathrm{ex}_M(m,n) \geq (\ell-1)m + (k-1)n - (k-1)(\ell-1).$$

Thus, for every k-permutation matrix P, $c_P \geq 2(k-1)$.

Proof. Take any 1-entry $M_{\alpha,\beta}$ of M. Let A be the $m \times n$ $\{0,1\}$-matrix with

$$a_{i,j} = \begin{cases} 0 & \text{if } \alpha \leq i \leq m-k+\alpha \text{ and } \beta \leq j \leq n-k+\beta \\ 1 & \text{otherwise.} \end{cases}$$

Then $M_{\alpha,\beta}$ cannot be represented by any 1-entry of A and so A avoids M. The number of 1-entries in A is

$$mn - (m-k+1)(n-\ell+1) = (\ell-1)m + (k-1)n - (k-1)(\ell-1).$$

□

We say that a $k \times \ell$ matrix M is *minimalist* if

$$\mathrm{ex}_M(m,n) = (\ell-1)m + (k-1)n - (k-1)(\ell-1).$$

The $k \times k$ *identity matrix* I_k is the permutation matrix with 1-entries on positions (i,i). As a consequence of a result of Füredi and Hajnal (see Corollary 1.9), every identity matrix I_k is minimalist.

In Section 1.2, we show that with one exception, all linear forbidden matrices with at most four 1's and no empty row or column are minimalist. We then show that adding an empty row to a minimalist matrix with 4 1-entries can create a matrix that is not minimalist.

The first example of a linear forbidden matrix that is not minimalist was constructed by the author [Cib09] and is presented in Section 1.4. The construction shows that for every k, $c_k \geq (k-1)^2/2$.

Theorem 1.3. *For every even $k \geq 2$, there exists a $k \times k$ permutation matrix C satisfying $c_C \geq \frac{k^2}{2}$ and so for every $k \geq 1$,*

$$c_k \geq (k-1)^2/2.$$

A *layered permutation* is a concatenation of decreasing sequences, called layers, such that each entry of one layer is smaller than entries of the following layers. Let $\rho(\ell_1, \ldots, \ell_q)$ be the layered permutation with m layers of lengths ℓ_i. A *layered matrix* $L(\ell_1, \ldots, \ell_q)$ is a permutation matrix corresponding to the layered permutation $\rho(\ell_1, \ldots, \ell_m)$. For example, $\rho(1,2,1)$ is the permutation 1324 and its permutation matrix $L(1,2,1)$ is M_3 from Section 1.2. Another example is the identity matrix corresponding to the layered permutation with all layers of length 1.

Consider matrices M of size $k_1 \times \ell_1$ and N of size $k_2 \times \ell_2$. The *direct sum* $M \oplus N$ of M and N is the $(k_1+k_2) \times (\ell_1+\ell_2)$ matrix with M on the intersection of the top k_1 rows and left ℓ_1 columns and N on the intersection of the bottom k_2 rows and right ℓ_2 columns and with 0's everywhere else. That is,

$$M \oplus N = \begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix}.$$

The *reversal matrix* J_k is the k-permutation matrix with 1-entries only on the main skew diagonal, that is, the entry at (i, j) is 1 if and only if $i + j = k + 1$. Notice that a layered matrix $L(\ell_1, \ldots, \ell_q)$ is a direct sum of q reversal matrices.

Backelin, West and Xin [BWX07] showed that for every k and n and every permutation matrix P, the n-permutation matrices avoiding $I_k \oplus P$ are in bijection with the n-permutation matrices avoiding $J_k \oplus P$. The particular cases of $k = 2$ and $P = I_1$ [SS85], $k = 2$ [Wes90] and $k = 3$ [BW00] were shown earlier. De Mier [dM07] generalized this to all $\{0, 1\}$-matrices.

Theorem 1.4 (de Mier, 2007)**.** *For every integers k, m and n and every matrix M, the set of $m \times n$ matrices avoiding $I_k \oplus M$ is in bijection with the set of $m \times n$ matrices avoiding $J_k \oplus M$. Moreover, the bijection preserves the number of 1-entries in every row and every column.*

In particular, for an arbitrary matrix M, the extremal functions of $I_k \oplus M$ and $J_k \oplus M$ are the same.

Corollary 1.5. *For every t, u and v, the matrix $J_t \oplus I_u \oplus J_v$ is minimalist. In particular, every layered matrix with at most two layers is minimalist.*

Let $\lambda_{\ell_1,\ldots,\ell_q}(m, n)$ be the extremal function of $L(\ell_1, \ldots, \ell_q)$. In Section 1.3, we give an upper bound on the extremal function of layered matrices. For every layered matrix L, we construct L-avoiding square matrices showing that the bound is tight up to the smaller order terms.

Theorem 1.6. *For every $q \geq 1$ and every q-tuple of layer lengths $\ell_i \geq 1$, the extremal function of the layered matrix $L(\ell_1, \ldots, \ell_q)$ satisfies*

$$\lambda(\ell_1, \ldots, \ell_q)(n, n) \geq \left(\ell_1 + \ell_q - q + 1 + 2\sum_{i=2}^{q-1} \ell_i\right) 2n + O(n^{1/2}),$$
$$\lambda(\ell_1, \ldots, \ell_q)(m, n) \leq \left(\ell_1 + \ell_q - q + 1 + 2\sum_{i=2}^{q-1} \ell_i\right) (m + n).$$

This is in contrast with the situation in the area of counting permutations avoiding a layered permutation (see Section 2.1), where the Stanley–Wilf limit is not known even for the layered permutation 1324.

We conclude with the description of all known linear matrices.

The following forbidden matrices are known to be linear:

- Double permutation matrices [Gen09]. A *double permutation matrices* is a matrix formed from a permutation matrix by doubling every column. See Fig. 1.1 b) for an example.
- Matrices from Fig. 1.1 a).

Let M and M' be linear forbidden matrices. The following operations preserve linearity of the forbidden matrix.

- Rotation and mirroring.
- Removing 1-entries.

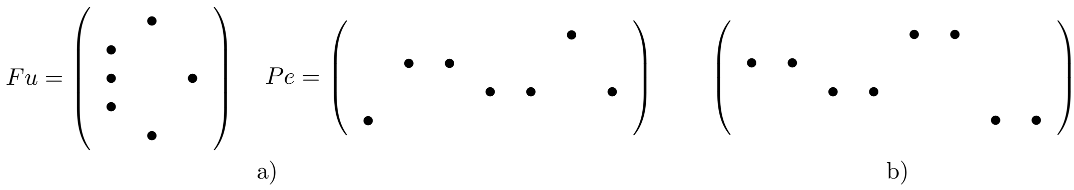


Figure 1.1: a) Linear forbidden matrices. The linearity of Fu was shown by Fulek [Ful09] and of Pe by Pettie [Pet11c]. b) An example of a double permutation matrix.

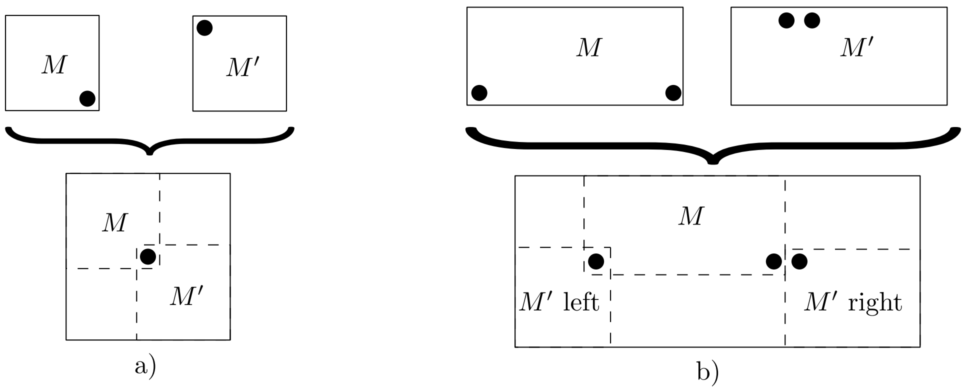


Figure 1.2: Operations producing linear forbidden matrices. a) Keszegh's composition of matrices. b) Pettie's grafting.

- Removing and adding empty rows and columns [Tar05]. See also Lemma 1.15.
- Adding a new row or column with a single 1-entry to the boundary of M so that the new 1-entry is next to an existing one [FH92]. See also Lemma 1.8.
- Adding a new row with a single 1-entry between two rows of M so that the new 1-entry is next to two existing ones [Tar05]. See also Lemma 1.14.
- Merging M and M' by identifying the bottom right corner of M and the top left corner of M' if the two identified corners contain 1-entries [Kes09]. See Fig. 1.2 a) for an example.
- A matrix is *ascending* if for every 1-entry (i, j), there is no 1-entry on the intersection of rows sharply below i and columns sharply to the left of j. A matrix is *descending* if its image under mirroring by vertical axis is ascending. Let M have all intervals of at least two rows non-ascending or all non-descending. Let M' have a pair of neighboring 1's in the top row and let M have 1's in the bottom left and the bottom right corner. Split M' into left and right part by a cut in between the rows containing the two neighboring 1's in the top row. Identify the top right corner of the left part with the bottom left corner of M and place the right part so that its top left corner is one column to the right and on the same row as the bottom right corner of M. This operation is called *grafting* M on M' and the resulting matrix is linear [Pet11c]. See Fig. 1.2 b) for an example.

- If M is of the form $R \oplus S$, M' is of the form $S \oplus T$ and S is nonempty, then $R \oplus S \oplus T$ is linear by Claim 1.18.

1.2 Matrices with four 1-entries

Definition 1.7. *Let P and P' be matrices such that P' can be created from P by adding a new row (or column) as the new first or last row (column) and this new row (column) contains a single* 1*-entry next to a* 1*-entry of P. We say that P' was created by an* elementary operation *from P.*

The following lemma and its corollary are a slightly modified version of a result of Füredi and Hajnal [FH92, Theorem 2.2].

Lemma 1.8 (Füredi and Hajnal, 1992)**.**

1. *If a matrix M contains a matrix M' then $\mathrm{ex}_{M'}(m,n) \leq \mathrm{ex}_M(m,n)$.*
2. *If a matrix M' is created by an elementary operation from a $k \times \ell$ matrix M, then for all $m, n \geq 0$*

$$\begin{aligned} \mathrm{ex}_{M'}(m,n) &\leq \mathrm{ex}_M(m,n) + m - k + 1 && \text{if we added a column, and} \\ \mathrm{ex}_{M'}(m,n) &\leq \mathrm{ex}_M(m,n) + n - \ell + 1 && \text{if we added a row.} \end{aligned}$$

Proof. The first part follows from the observation that if a matrix A avoids M' then it avoids M.

In the second part, by symmetry, it suffices to consider the case when the last column was added. Let i and $\ell + 1$ be the row and column of the new 1-entry in M'. Assume that a matrix A' avoids M'. By removing from A' the rightmost 1-entry in each of the rows number i up to $m-k+i$, we obtain a matrix avoiding M with at least $\mathrm{ex}_{M'}(m,n) - m + k - 1$ ones. □

A matrix M is *tightly contained* in a matrix T if M is contained in T and the sizes of M and T are equal.

Corollary 1.9 (Füredi and Hajnal, 1992)**.** *Let M be a $k \times \ell$ minimalist matrix. Let M' be a $k' \times \ell'$ nonempty matrix tightly contained in a matrix T obtained from M by applying several elementary operations. Then M' is minimalist.*

Proof. By Claim 1.2, the extremal function of M' cannot be smaller than the extremal function of a minimalist matrix of the same size.

Thus it is enough to show that applying a single elementary operation on a minimalist matrix M produces a minimalist matrix T. By symmetry, it is enough to consider an elementary operation that adds a new column.

We obtain

$$\begin{aligned} \mathrm{ex}_T(m,n) &\leq \mathrm{ex}_M(m,n) + m - k + 1 \\ &= (\ell - 1)m + (k-1)n - (k-1)(\ell-1) + m - k + 1 \\ &= \ell m + (k-1)n - (k-1)\ell. \end{aligned}$$

□

We define the following matrices: $M_0 = \begin{pmatrix} \bullet \end{pmatrix}$,

$$M_1 = \begin{pmatrix} & \bullet & \\ \bullet & & \\ \bullet & & \bullet \end{pmatrix}, \; M_2 = \begin{pmatrix} & \bullet \\ \bullet & \\ \bullet & \\ & \bullet \end{pmatrix}, \; M_2' = \begin{pmatrix} & \bullet \\ \bullet & \\ \bullet & \bullet \end{pmatrix} \text{ and } M_3 = \begin{pmatrix} \bullet & & & \\ & & \bullet & \\ & \bullet & & \\ & & & \bullet \end{pmatrix}.$$

Füredi and Hajnal [FH92] started and Tardos [Tar05] finished the characterization of the order of magnitude of $\mathrm{ex}_n(P)$ for all P with at most four 1-entries.

Let $\mathcal{L}_4$ be the set of linear forbidden matrices with four 1-entries and with at least one 1-entry in every row and in every column.

Every matrix in $\mathcal{L}_4$ other than M_3 is tightly contained in some matrix that can be obtained from M_0, M_1 or M_2 by a sequence of elementary operations. Apparently, M_0 is minimalist.

The following definition will be useful in proving that M_1 and M_2 are minimalist.

Definition 1.10. *Let $t_A(j)$ be the smallest (topmost) row t such that the matrix A contains a 1-entry on position (t, j). Let $b_A(j)$ be the largest (lowest) row b such that A contains a 1-entry on position (b, j).*

Let A be a matrix and B a matrix tightly contained in A. We say that the 1-entry (i', j') of A finds a 1-entry (i, j) of B if j is the largest column smaller than j' satisfying $t_A(j) \leq i' < b_B(j)$ and i is the smallest row larger than i' such that (i, j) is a 1-entry of B. See Fig. 1.3 b).

Lemma 1.11 (Tardos, 2005)**.** *The matrix M_1 is minimalist.*

Proof. Let A be an $m \times n$ matrix avoiding M_1 and let B be the matrix obtained from A by removing the topmost 1-entry in every column. We now show that every 1-entry of A falls within one of the following categories.

(i) Topmost 1-entry in some column of A.

(ii) One of the two rightmost 1-entries in one of the rows $2, \ldots, m$ of B.

(iii) A 1-entry of B found by $(t_A(j'), j')$, where $j' \in \{2, \ldots, n-1\}$.

For a contradiction, assume that a 1-entry (i, j) of A is in none of the categories. Let j', j'' be the columns of some two 1-entries of B in row i to the right of (i, j). We can find such j', j'' since (i, j) is not of type (ii). Because each of (i, j), (i, j') and (i, j'') is in B, they are not the topmost entries in their columns. Let r be the row of the nearest 1-entry above (i, j) in column j.

If some column $u \in \{j+1, \ldots, j'\}$ satisfies $t_A(u) < r$ then the intersection of rows $t_A(u), r, i$ and columns j, u, j'' contains an occurrence of M_1. See Fig. 1.3 a).

Otherwise, we take the smallest $u' > j$ such that $r \leq t_A(u') < i$. There is such $u' \leq j'$ since $r \leq t_A(j') < i$. Then $(t(u'), u')$ finds (i, j) and $u' \in \{2, \ldots, n-1\}$, a contradiction. See Fig. 1.3 b).

The matrix A contains n 1-entries of type (i), $2m - 2$ of type (ii) and $n - 2$ of type (iii). $\square$

A small modification of the proof can be used to improve the linear upper bound on the extremal function of M_2 found by Füredi and Hajnal [FH92].

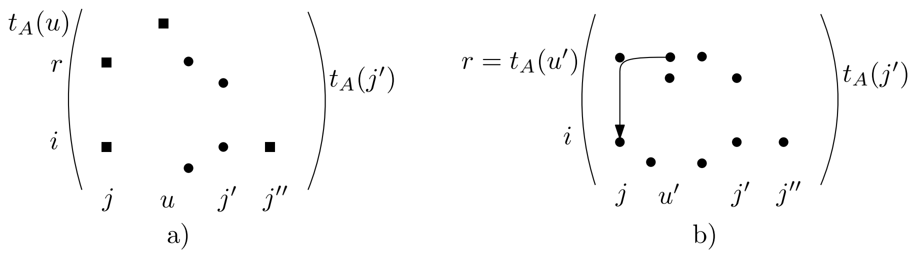


Figure 1.3: Proof of Lemma 1.11. Arrow represents finding of a 1-entry. Squares represent an occurrence of M_1. Row indices increase from top to bottom and column indices from left to right.

Lemma 1.12. *The matrix M_2 is minimalist.*

Proof. Let A be an $m \times n$ matrix avoiding M_2 and let B be the matrix obtained from A by removing both the topmost and the lowermost 1-entry in every column.

We now show that every 1-entry falls within one of the following categories. This will imply that A has at most $m + 3n - 3$ 1-entries.

(i) Topmost or lowermost 1-entry in some column of A.

(ii) The rightmost 1-entry in one of the rows $2, \ldots, m-1$ of B.

(iii) A 1-entry of B found by $(t_A(j'), j')$, where $j' \geq 2$.

For contradiction, assume that a 1-entry (i, j) of A is in none of the categories. Let j' be the column of some 1-entry of B in row i to the right of (i, j). Let r be the row of the nearest 1-entry above (i, j) in column j.

If some column $u \in \{j+1, \ldots, j'\}$ satisfies $t_A(u) < r$ and $b_A(r) > i$ then the intersection of rows $t_A(u), r, i, b_A(u)$ and columns j, u contains an occurrence of M_2. See Fig. 1.4 a).

Otherwise, in particular $t_A(j') \geq r$ since $b_A(j') > i$. We take the smallest $u' > j$ such that $r \leq t_A(u') < i < b_A(u')$. There is such $u' \leq j'$ since $r \leq t_A(j') < i < b_A(j')$.

If $(t(u'), u')$ finds some $(w, v) \neq (i, j)$, then $j < v < u'$. By the definition of finding, $t_A(v) \leq t_A(u')$. We also have $b_A(v) \geq b_A(u')$, since otherwise, the intersection of rows $t_A(u'), w, b_A(v), b_A(u')$ and columns v, u' contains M_2. See Fig. 1.4 b). But then v should have been selected instead of u'. Thus $(t(u'), u')$ finds (i, j) and since $u' \geq 2$, we get a contradiction.

□

Corollary 1.9 immediately implies the following corollary.

Corollary 1.13. *Every matrix $M \in \mathcal{L}_4$ different from M_3 is minimalist.*

For every $k \geq 4$ let $\hat{N}_k$ be the $k \times 2$ matrix formed from M_2 by adding $k-4$ empty rows after its second row. For $k \geq 3$ let $\hat{P}_k$ (N'_k) be the $k \times 3$ matrix formed from M_1 (M'_2) by adding $k-3$ empty rows after its second row. Let N_k (P_k) be the $k \times 2$ matrix formed from $\hat{N}_k$ ($\hat{P}_k$) by changing the entries in the left

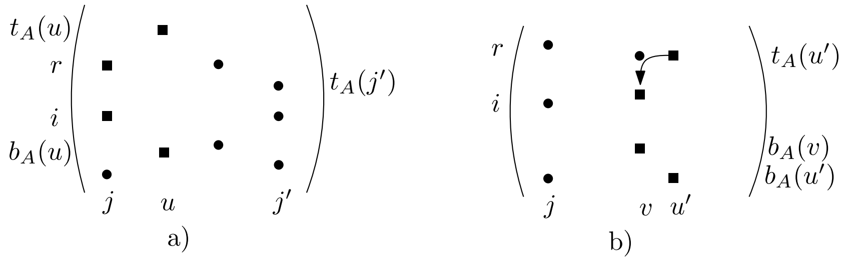
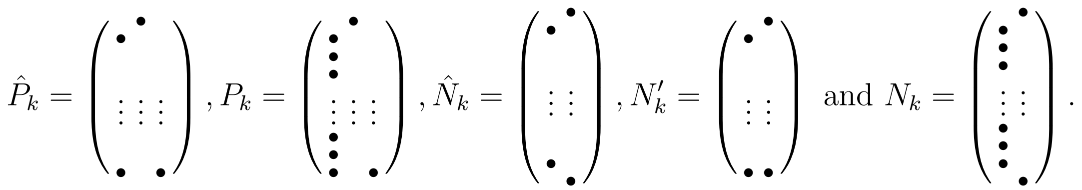


Figure 1.4: Proof of Lemma 1.12. Arrow represents finding of a 1-entry. Squares represent an occurrence of M_2.

column of the newly added rows to 1. That is, $N_4 = \hat{N}_4 = M_2$, $P_3 = \hat{P}_3 = M_1$, $N'_3 = M'_2$,



The following two lemmas are slight strengthenings of results of Tardos [Tar05, Lemma 2.3].

Lemma 1.14 (Tardos 2005)**.** *Let M be a $k \times l$ matrix. Let M' be a $k' \times l$ matrix formed from M by removing every row i containing a single 1-entry such that the rows $i-1$ and $i+1$ contain 1-entries in the same column. Let u be the smallest number such that from each sequence of u consecutive rows of M at least one remains as a row of M'. Then for every $m, n \geq 1$, $\mathrm{ex}_M(m,n) \leq u\,\mathrm{ex}_{M'}(m,n)$.*

If, in addition, every row of M contains at most one 1-entry, then we have $\mathrm{ex}_M(m,n) \leq u\,\mathrm{ex}_{M'}(\lceil m/u \rceil, n)$.

Proof. Notice that for every sequence of consecutive rows removed from M, all rows of the sequence have the single 1-entry in the same column.

Let A be an $m \times n$ M-avoiding matrix with $\mathrm{ex}_M(m,n)$ 1-entries. A 1-entry of A is *good* if it is $(uk+1)$-st 1-entry from top in its column for some nonnegative integer k. Let A' be the $m \times n$ matrix obtained from A by changing to 0 all 1-entries except the good ones. Then A' has at least $\mathrm{ex}_M(m,n)/u$ 1-entries and avoids M'.

To *contract* a u-tuple of rows means to replace them with a single row with a 1-entry in the columns where at least one of the original rows has a 1-entry. To prove the second part, we split all the rows of A' into $\lceil m/u \rceil$ groups of at most u consecutive rows. We create A'' by contracting these $\lceil m/u \rceil$ groups of rows of A'. If A'' contained M' then A' would contain M' as well, since no two 1-entries of M' lie in the same row. Thus A'' avoids M'. Since there are at least $u-1$ rows between two 1-entries in the same column of A', A'' has the same number of 1-entries as A' has. $\square$

Lemma 1.15 (Tardos 2005)**.** *Let M be a $k \times \ell$ matrix and let M' be the $k' \times \ell'$ matrix obtained from M by removal of all empty rows and columns. Let u and v be the smallest numbers such that each sequence of u consecutive columns and* *each sequence of v consecutive rows of M contains a 1-entry. Then for every $m, n \geq 1$*

$$\mathrm{ex}_M(m,n) \leq uv\,\mathrm{ex}_{M'}(\lceil m/v\rceil, \lceil n/u\rceil).$$

Proof. Let A be a matrix avoiding M. Let M'' be the matrix obtained from M by removal of all empty columns. For every i, let A_i, $i \in [u]$ be the matrix formed from all the columns j of A such that $j \equiv i \pmod u$. All A_i avoid M'', have at most $\lceil n/u\rceil$ columns and one of them has at least $\mathrm{ex}_M(m,n)/u$ 1-entries. Thus $\mathrm{ex}_M(m,n) \leq u\,\mathrm{ex}_{M''}(m, \lceil n/u\rceil)$ and applying the same argument to the rows gives $\mathrm{ex}_{M''}(m, \lceil n/u\rceil) \leq v\,\mathrm{ex}_{M'}(\lceil m/v\rceil, \lceil n/u\rceil)$. □

Füredi and Hajnal claim that their proof of $\mathrm{ex}_{M_2}(n) \leq 7n$ can be generalized to $\mathrm{ex}_{N_k} \leq (k+3)n$. The proof lacks details and the conclusion is wrong. After a correction, it gives the upper bound $\mathrm{ex}_{N_k} \leq (3k-5)n$. [1]

Lemmas 1.14, 1.11 and 1.12 imply the following upper bounds for P_k and N_k.

Corollary 1.16. *For every fixed $k \geq 3$,*

$$\mathrm{ex}_{\hat{P}_k}(m,n) \leq \mathrm{ex}_{P_k}(m,n) \leq 2m + (2k-4)n + O(1).$$

and for every fixed $k \geq 4$,

$$\mathrm{ex}_{\hat{N}_k}(m,n) \leq \mathrm{ex}_{N_k}(m,n) \leq m + (3k-9)n + O(1)$$

We show that $\hat{P}_k$ for $k \geq 4$ and $\hat{N}_k$ for $k \geq 5$ are not minimalist and that the upper bound on $\mathrm{ex}_{\hat{P}_k}(n)$ is essentially tight.

Theorem 1.17. *For every fixed $k \geq 3$ and $m \geq 2k(\sqrt{n}+1)$,*

$$\mathrm{ex}_{P_k}(m,n) \geq \mathrm{ex}_{\hat{P}_k}(m,n) \geq 2m + (2k-4)n - O(1+n/m)$$

and, when in addition $k \geq 4$,

$$\mathrm{ex}_{N_k}(m,n) \geq \mathrm{ex}_{\{\hat{N}_k, N'_{k-1}\}}(m,n) \geq m + (2k-5)n - O(1+n/m).$$

Proof. We first construct the $m \times n$ matrix A avoiding $\hat{P}_k$. Let $\mathcal{B}$ be the union of the bottom $b := k\lceil 2kn/(m-2k)\rceil$ rows and let $\mathcal{T}$ be the union of all the other rows. Since $(m-2k)^2 \geq 4k^2 n$, we have $2kn/(m-2k) \leq (m-2k)/(2k)$ and so $b \leq m/2$.

Let $w := \lfloor (m-b)/k \rfloor$ and let $c := \lceil (n-2)/w\rceil$. Thus $w \geq (m/2-k)/k$ and

$$kc \leq k\lceil nk/(m/2-k)\rceil \leq k\lceil 2nk/(m-2k)\rceil = b.$$

For every $i \in [c]$ let $\mathcal{C}_i$ be the set of w columns $(i-1)w+1, \ldots, \min\{iw, n-2\}$. That is, all columns except the last (rightmost) two are in some $\mathcal{C}_i$. The last two columns have 1-entries in the rows of $\mathcal{T}$ and 0's in the rows of $\mathcal{B}$. Every column of $\mathcal{C}_i$ has 1-entries in the $k-2$ rows $n - i(k-2)+1, \ldots, n-(i-1)(k-2)$ and no other 1-entries in the rows of $\mathcal{B}$. The j-th column of every $\mathcal{C}_i$ has 1-entries in

[1] The proof of the upper bound $\mathrm{ex}_{M_2}(n) \leq 7n$ can be generalized by changing the first category of 1-entries from "first or last in their row" to "one of the first $k-3$ or last $k-3$ in their row" and by allowing, for every α, up to $k-3$ 1-entries in the same row to have type α.

the $k-2$ rows $(j-1)(k-2)+1, \ldots, j(k-2)$ and no other 1-entries in the rows of $\mathcal{T}$. See Fig. 1.5(a).

Every column other than the last two contains $2k-4$ 1-entries and the last two columns contain 1-entries in all but the last b rows. The number of 1-entries in A is thus

$$(2k-4)(n-2)+2m-2b \geq (2k-4)n+2m-O(1+n/m).$$

Assume that A contains $\hat{P}_k$. Let c_1, c_2 and c_3 be the columns of the occurrence of $\hat{P}_k$ with $c_1 < c_2 < c_3$. Let α be the top 1-entry in the leftmost column of $\hat{P}_k$, β the bottom 1-entry in the leftmost column, γ the 1-entry in the middle column and δ the 1-entry in the rightmost column.

All columns except the last two have all 1-entries in $\mathcal{T}$ within an interval of length $k-2$. Because α and β must have $k-3$ rows in between, the occurrence of $\hat{P}_k$ does not lie entirely on the rows of $\mathcal{T}$. Consequently, β and δ are be mapped inside $\mathcal{B}$ and c_3 is not one of the last two columns.

Since β and δ are in the same row, all c_1, c_2 and c_3 are in the same set $\mathcal{C}_i$. Let r be the row containing β and δ. Since α and γ are at least $k-2$ rows above r, they are in $\mathcal{T}$. But $c_1, c_2 \in \mathcal{C}_i$ and since $c_1 < c_2$, every 1-entry in $\mathcal{C}_i \cap c_1$ lies above the 1-entries in $\mathcal{C}_i \cap c_2$, a contradiction.

The construction of the $m \times n$ matrix D avoiding both $\hat{N}_k$ and N'_{k-1} shares many similarities with the construction of A. The values b, w and c and the sets $\mathcal{B}$, $\mathcal{T}$ and $\mathcal{C}_i$ are defined in the same way.

The last column has 1-entries in the rows of $\mathcal{T}$ and 0's in the rows of $\mathcal{B}$. The last but one column has no 1-entry. Every column of $\mathcal{C}_i$ has 1-entries in the $k-2$ rows $n-i(k-2)+1, \ldots, n-(i-1)(k-2)$ and no other 1-entries in the rows of $\mathcal{B}$. The j-th column of every $\mathcal{C}_i$ has 1-entries in the $k-3$ rows $(j-1)(k-3)+1$, $\ldots$, $j(k-3)$ and no other 1-entries in the rows of $\mathcal{T}$. See Fig. 1.5(b).

The number of 1-entries in D is

$$(2k-5)(n-2)+(m-b) \geq (2k-5)n+2m-O(n/m+1).$$

Assume that D contains $\hat{N}_k$ or N'_{k-1}. Let c_1 and c_2 be the columns of the occurrence and let $c_1 < c_2$. Let α and β be the top and bottom 1-entry in the left column of the contained matrix, and let γ and δ the top and bottom 1-entry in the right column.

All columns of D except the last one have all 1-entries in $\mathcal{T}$ within an interval of length $k-3$. Because α and β must have $k-4$ rows in between, β is mapped on a row of $\mathcal{B}$. Consequently, δ is mapped on a row of $\mathcal{B}$ and c_2 cannot be one of the last two columns.

Since δ is to the right and below or at the same row as β, both c_1 and c_2 are in the same set $\mathcal{C}_i$. Let r be the row containing β. Since α and γ are at least $k-3$ rows above r, they are in $\mathcal{T}$. But $c_1, c_2 \in \mathcal{C}_i$ and since $c_1 < c_2$, every 1-entry in $\mathcal{C}_i \cap c_1$ lies above the 1-entries in $\mathcal{C}_i \cap c_2$, a contradiction.

□

1.3 Layered patterns

The following claim and its proof is an analogue of a result of Claesson, Jelínek and Steingrímsson [CJS12] from the realm of counting permutations with forbidden

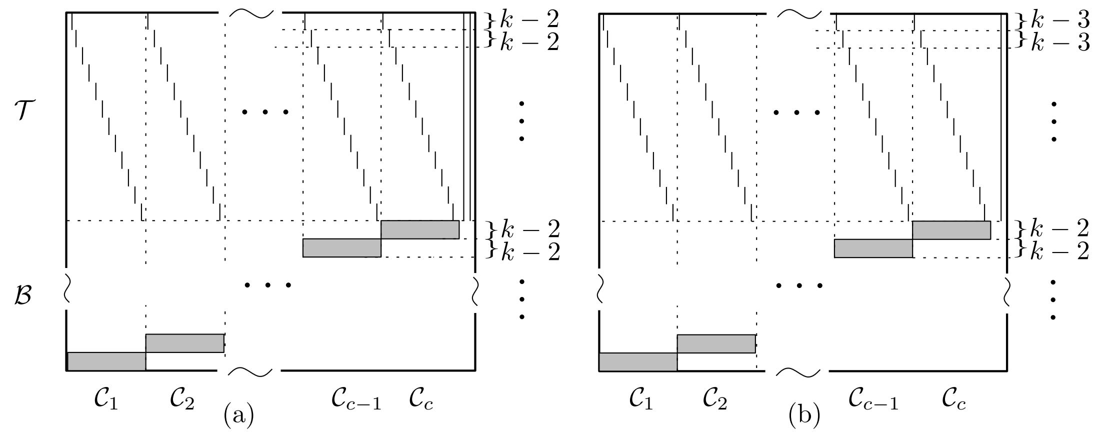


Figure 1.5: The gray areas and the full lines (other than the boundary) represent 1-entries. (a) The matrix $A_k(m,n)$ avoiding $\hat{P}_k$. (b) The matrix $D_k(m,n)$ avoiding $\hat{N}_k$ and N'_{k-1}.

patterns.

Claim 1.18. *Let R, S and T be matrices such that the first and the last row and the first and the last column of S is nonempty. Then $\mathrm{ex}_{R\oplus S\oplus T}(m,n) \leq \mathrm{ex}_{R\oplus S}(m,n) + \mathrm{ex}_{S\oplus T}(m,n)$.*

Proof. Let A be a matrix avoiding $R \oplus S \oplus T$.

Start with all 1-entries of A uncolored. Take columns of A from left to right. For every column j, take 1-entries in column j from top to bottom. Color the 1-entry $A_{i,j}$ red if if coloring it blue would form a blue occurrence of $R \oplus S$ or if there is a red 1-entry $A_{i',j'}$ to the northwest of (i,j), that is, with $i' < i$ and $j' < j$. Color $A_{i,j}$ blue otherwise.

Let the *last* 1*-entry of a matrix* M be the 1-entry from the rightmost nonempty column of M with the largest row index. Notice that for every red 1-entry (i,j) of A, there is an occurrence of $R \oplus S$ with the last 1-entry of $R \oplus S$ mapped to some 1-entry (i',j') with $i' \leq i$ and $j' \leq j$.

By the definition, the blue entries avoid $R \oplus S$, thus A has at most $\mathrm{ex}_{R\oplus S}(m,n)$ blue 1's.

The claim will be proven when we show that A contains no $S \oplus T$ with all 1-entries red. For a contradiction, we assume that A contains a red $S \oplus T$ and fix one of its occurrences. The occurrence of S (T) within this occurrence of a red $S \oplus T$ is referred to as the fixed occurrence of red S (T).

A *left-to-right* minimum in a matrix M is a 1-entry (i,j) that is the only 1-entry in the submatrix of M formed by its first i rows and j columns. Let $s_1 \dots s_\alpha$ be the left-to-right minima of S and let $(r_1,c_1) \dots (r_\alpha,c_\alpha)$ be the positions of $s_1 \dots s_\alpha$ in the fixed occurrence of red S. See Fig. 1.6.

Since (r_α,c_α) is red, we can fix a red 1-entry $(\bar{r},\bar{c})$ satisfying $\bar{r} \leq r_\alpha$ and $\bar{c} \leq c_\alpha$ and an occurrence of $R \oplus S$ satisfying the following. All 1-entries of the occurrence of the fixed $R \oplus S$ are blue except for one 1-entry from the last column of $R \oplus S$ that is mapped to $(\bar{r},\bar{c})$. The occurrence of R within this occurrence of $R \oplus S$ is referred to as the fixed occurrence of R. The fixed occurrence of R is contained in the intersection of the first $r_\alpha - 1$ rows and $c_\alpha - 1$ columns of A (this

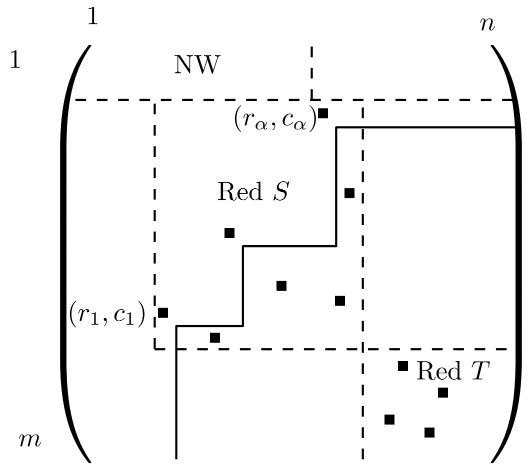


Figure 1.6: Proof of Claim 1.18. Squares represent red 1-entries of the occurrence of $S \oplus T$. All 1-entries below the full line are red. The northwest area (NW) contains a blue R.

area is marked "NW" in Fig. 1.6). The fixed occurrence of R is not contained within the first $c_1 - 1$ columns of A, since otherwise, the fixed occurrences of R, S and T would form an occurrence of $R \oplus S \oplus T$.

Since some column of the fixed occurrence of R is the column c_1 or is to the right of it, all the 1-entries of the fixed occurrence of $R \oplus S$ corresponding to the 1-entries of S are to the right of the column c_1. All blue 1's of A to the right of column c_1 are in the first r_1 rows, since otherwise, they would be red because of the red (r_1, c_1). Thus the whole fixed occurrence of $R \oplus S$ is in the intersection of the first r_1 rows and c_α columns of A and, together with the fixed occurrence of T, it forms $R \oplus S \oplus T$. □

We now finish the proof of the upper bound of Theorem 1.6.

Claim 1.19. *For every $q \geq 2$ and every q-tuple of layer lengths $\ell_i \geq 1$, the extremal function of the layered matrix $L(\ell_1, \dots, \ell_q)$ satisfies*

$$\lambda(\ell_1, \dots, \ell_q)(m, n) \leq \left(\ell_1 + \ell_q - q + 1 + 2 \sum_{i=2}^{q-1} \ell_i \right) (m + n).$$

Proof. We proceed by induction on q. Corollary 1.5 proves the claim for $q = 2$. Let $L = L(\ell_1, \dots, \ell_q)$ be a layered matrix with $q \geq 3$ layers. The matrix L is a direct sum $L = L_1 \oplus L_2 \oplus L_3$ of $L_1 = L(\ell_1, \dots, \ell_{q-2})$, $L_2 = L(\ell_{q-1})$ and $L_3 = L(\ell_q)$. By Claim 1.18,

$$\begin{aligned}
\mathrm{ex}_L(m, n) &\leq \mathrm{ex}_{L_1 \oplus L_2}(m, n) + \mathrm{ex}_{L_2 \oplus L_3}(m, n) \\
&\leq \left(\ell_1 + \ell_{q-1} - q + 2 + 2 \sum_{i=2}^{q-2} \ell_i + \ell_{q-1} + \ell_q - 1 \right) (m + n) \\
&= \left(\ell_1 + \ell_q - q + 1 + 2 \sum_{i=2}^{q-1} \ell_i \right) (m + n).
\end{aligned}$$

□

We now construct square matrices attaining the lower bound of Theorem 1.6.

Lemma 1.20. *For every layered matrix $L = L(\ell_1, \dots, \ell_q)$ there exists an $n \times n$ matrix A avoiding L and having at least the following number of* 1*-entries.*

$$\lambda(\ell_1, \dots, \ell_q)(n, n) \geq \left(2\ell_1 + 2\ell_q - 2q + 2 + 4\sum_{i=2}^{q-1} \ell_i\right) n - O(n^{1/2}).$$

Proof. Let (a, b) be a pair of integers. Let the $[a, b]$-*strip* $S[a, b]$ be the set of elements of A at positions (i, j) satisfying $j - i \in [a, b]$. Let the $[a, b]$-*antistrip* $\bar{S}[a, b]$ be the set of elements of A at positions (i, j) satisfying $i + j \in [a, b]$. The *width* of a strip (antistrip) is the difference $b - a + 1$.

First we set to 1 all the elements of $q - 1$ antistrips of A placed near the main skew diagonal and with distances between two consecutive antistrips roughly $\sqrt{n}$. The antistrips are $S_i := \bar{S}[a_i, b_i]$, where $i \in [q - 1]$ and

$$\begin{aligned} &a_1 = n - \frac{q}{2}\sqrt{n}, \quad a_i = b_{i-1} + \sqrt{n} + 1 \text{ for } i \in [2, q-1] \text{ and} \\ &b_i = a_i + \ell_i + \ell_{i+1} - 1 \text{ for } i \in [q-1]. \end{aligned}$$

The number of 1-entries in the i-th antistrip is

$$(\ell_i + \ell_{i+1})n - O(n^{1/2}).$$

For $i = 1, \dots, q$, we let the *i-th belt* B_i be the antistrip of width $\sqrt{n}$ between the antistrips S_i and S_{i+1} (if $i \in [2, q-1]$), to the left of the S_1 ($i = 1$), or to the right of S_{q-1} ($i = q$). That is,

$$\begin{aligned} B_1 &= S[a_1 - \sqrt{n}, a_1 - 1], \\ B_i &= S[b_i + 1, a_{i+1} - 1] \qquad \text{when } 2 \leq i \leq q-2 \text{ and} \\ B_{q-1} &= S[b_{q-1} + 1, b_{q-1} + \sqrt{n}]. \end{aligned}$$

For every $i = 1, \dots, q$, we put 1 in the elements of the set $\mathcal{B}_i$, which is the union of the intersections $B_{i,j} = B_i \cap S[r_{i,j}, s_{i,j}]$. If $i = 1$ or $i = q$ and $j \in [1, \sqrt{n} - 1]$, the strip $S[r_{i,j}, s_{i,j}]$ is defined by

$$\begin{aligned} &r_{i,1} = -n, \qquad r_{i,j} = s_{i,j-1} + 2(\sqrt{n} + \ell_i) \text{ for } j \in [2, \sqrt{n} - 1] \text{ and} \\ &s_{i,j} = r_{i,j} + 2\ell_i - 2 \text{ for } j \in [\sqrt{n} - 1]. \end{aligned}$$

For $i \in \{2, \dots, q-1\}$, we have

$$\begin{aligned} &r_{i,1} = -n, \quad r_{i,j} = s_{i,j-1} + \sqrt{n} + 2\ell_i \text{ for } j \in [2, 2\sqrt{n} - 1] \text{ and} \\ &s_{i,j} = r_{i,j} + 2\ell_i - 2 \text{ for } j \in [2\sqrt{n} - 1]. \end{aligned}$$

See Fig. 1.7.

All belts B_i are at distance $O(\sqrt{n})$ from the main skew diagonal, where the constant hidden by the O-notation depends on q. Thus for every i, all except for

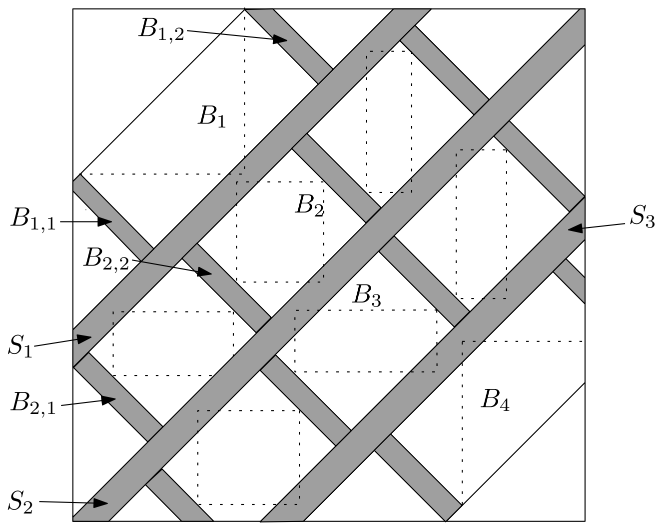


Figure 1.7: The belts B_i and antistrips S_i. Grey areas represent 1-entries.

$O(1)$ of the sets $B_{i,j}$ contain $(\ell_i - 1)\sqrt{n}$ 1-entries. The total number of 1-entries in B_i is then

$$(\ell_i - 1)n - O(\sqrt{n}) \qquad \text{when } i \in \{1, q\} \text{ and}$$
$$(2\ell_i - 2)n - O(\sqrt{n}) \qquad \text{when } i \in [2, q-1].$$

Let $\mathcal{B}_i^+ = \bigcup_{j=i}^{q} \mathcal{B}_j \bigcup_{j=i}^{q-1} S_j$ and let $S_i^+ = \bigcup_{j=i+1}^{q} \mathcal{B}_j \bigcup_{j=i}^{q-1} S_j$

The *distance between entries* $\alpha = (r, c)$ and $\beta = (r', c')$ is the L_1-distance, that is, $\mathrm{dist}(\alpha, \beta) = |r - r'| + |c - c'|$. We fix an occurrence of L in A and for every element α of L, let $\phi(\alpha)$ be the 1-entry of A representing α.

Claim 1.21. *Let α and γ be 1-entries of L from layers i' and i'' and let β and β' be 1-entries of L from layer i, where $i' < i < i''$. Then $\mathrm{dist}(\phi(\alpha), \phi(\gamma)) > \mathrm{dist}(\phi(\beta), \phi(\beta'))$.*

Proof. Since β and β' lie inside the axis-parallel rectangle with α and γ in its corners, the row and column differences have larger absolute values for $\phi(\alpha)$ and $\phi(\gamma)$ than for $\phi(\beta)$ and $\phi(\beta')$. □

We say that 1-entries $\alpha = (r, c)$, $\beta = (r', c')$ are in the *NW-SE direction* if the differences $r - r'$ and $c - c'$ are nonzero and have the same signs. The 1-entries α, β are in the *NE-SW direction* if $r - r'$ and $c - c'$ are nonzero and have different signs. If two 1-entries of a strip of width d are in the NW-SE direction, then their distance is at most $d - 1$.

Claim 1.22. *If two 1-entries of L are at distance d in the NE-SW direction (NW-SE direction) then they cannot be mapped to the same strip (antistrip) of width d. In particular:*

- *For every i, every 1-entry α from layer i and β from layer $i+1$ are in NW-SE direction and $\mathrm{dist}(\alpha, \beta) = \ell_i + \ell_{i+1}$. Thus, if $\phi(\alpha)$ is in the antistrip S_i then $\phi(\beta)$ cannot be in S_i and so it is in $\mathcal{B}_{i+1}^+$.*

- *The i-th layer contains two 1-entries at distance $2\ell_i - 2$ in the NE-SW direction and so for every j, not all 1-entries of the i-th layer are mapped to the strip $B_{i,j}$.*

To prove that A avoids L, we proceed in the following way. We say that a k-tuple of elements of A forms a *string of length k in S_i* if they all are elements of S_i of the form (r, c), $(r-1, c+1)$, $\ldots (r-k+1, c+k-1)$ for some pair of indices (r, c). We show that if there is an occurrence of L in A, we can, for every i, "push" the 1-entries of the i-th layer to a string of length ℓ_i in S_i. But then the 1-entries of the q-th layer remain unmapped.

Let i be the smallest index such that the 1-entries of the i-th layer are not mapped to a string in S_i. We aim on finding an occurrence of L in A, where each 1-entry α of L from a layer other than i is mapped to $\phi'(\alpha) = \phi(\alpha)$ and the 1-entries of the i-th layer are mapped to a string in S_i. If $i > 1$ then all 1-entries of layer $i-1$ are mapped to S_{i-1} and by Claim 1.22, the 1-entries of the i-th layer are in $\mathcal{B}_i^+$. For $i = 1$, all 1-entries of the first layer are in $\mathcal{B}_1^+$ trivially.

If $i = 1$, we let $r_0 = 0$ and $c_0 = 0$. Otherwise, we pick r_0 and c_0 so that the 1-entries of the $(i-1)$-st layer are mapped on the string of entries $(r_0, c_0 - \ell_{i-1} + 1) \ldots (r_0 - \ell_{i-1} + 1, c_0)$. Let α and ω be the southwesternmost and the northeasternmost entries of the i-th layer. Let $(r_\alpha, c_\alpha) = \phi(\alpha)$ and $(r_\omega, c_\omega) = \phi(\omega)$. Let R be the intersection of rows $[r_0+1, r_\alpha]$ and columns $[c_0+1, c_\omega]$. Then for every occurrence of J_{ℓ_i} in R, there is an occurrence of L in A that maps every 1-entry α in a layer $i' \neq i$ to $\phi(\alpha)$ and the 1-entries in layer i to the 1-entries of the occurrence of J_{ℓ_i}. Thus, it is enough to show that the rectangle R contains a string of length ℓ_i in S_i. This occurs when $r_\alpha + c_\omega \geq a_i + \ell_i - 1$. In such a case, we say that R is *sufficient*. See Fig. 1.8.

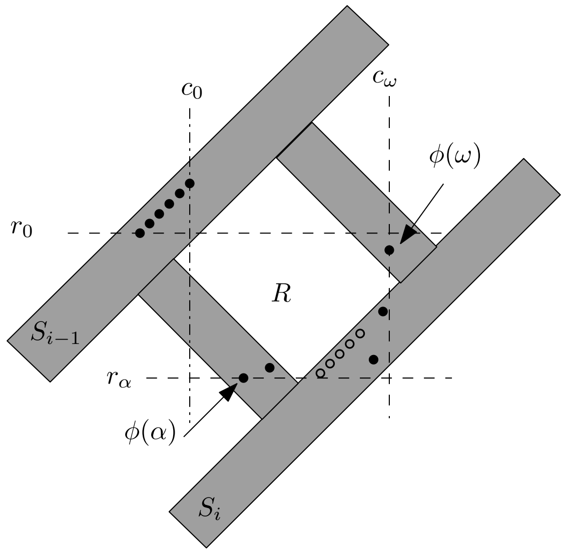


Figure 1.8: "Pushing" i-th layer to S_i. Full dots represent the mapping ϕ of 1's from the $(i-1)$-st and i-th layer. Empty circles are locations of the 1's of the i-th layer in the new mapping ϕ'.

First, we assume that some 1-entry β of the i-th layer is mapped to $\phi(\beta) \in S_i^+$, that is, $r + c \geq a_i$, where $(r, c) = \phi(\beta)$. Then for some $t \in [\ell_i]$, β is the t-th southwestern entry of the i-th layer. Then

$$r_\alpha \geq r + t - 1 \quad \text{and} \quad c_\omega \geq c + \ell_i - t.$$

Then the rectangle R is sufficient since $r_\alpha + c_\omega \geq r + c + \ell_i - 1 \geq a_i + \ell_i - 1$.

Otherwise, all 1-entries of the i-th layer are mapped to 1-entries in $\mathcal{B}_i$. Let j and j' be the indices such that $\phi(\alpha) \in B_{i,j}$ and $\phi(\omega) \in B_{i,j'}$. By Claim 1.22, $j \neq j'$. Thus, if $i = 1$,

$$r_\alpha - c_\alpha - (r_\omega - c_\omega) \geq r_{1,j+1} - s_{1,j} \geq 2(\sqrt{n} + \ell_1).$$

Since both $\phi(\alpha)$ and $\phi(\omega)$ are in $\mathcal{B}_1$ we have

$$r_\alpha + c_\alpha + r_\omega + c_\omega \geq 2(a_1 - \sqrt{n}).$$

Summing the two inequalities gives

$$2(r_\alpha + c_\omega) \geq 2(\sqrt{n} + \ell_1) + 2(a_1 - \sqrt{n}) = 2(a_1 + \ell_1).$$

and so the rectangle R is sufficient.

If $i > 1$,

$$r_\alpha - c_\alpha - (r_\omega - c_\omega) \geq r_{i,j+1} - s_{i,j} \geq \sqrt{n} + 2\ell_i$$

and since $c_\alpha \geq c_0 + 1$ and $r_\omega \geq r_0 + 1$,

$$r_\alpha + c_\omega \geq \sqrt{n} + 2\ell_i + r_0 + c_0 + 2.$$

Since the ℓ_{i-1} 1-entries of the $(i-1)$-st layer are mapped to a string in S_{i-1}, $r_0 + c_0 \geq a_{i-1} + \ell_{i-1} - 1$ and so

$$r_\alpha + c_\omega \geq \sqrt{n} + 2\ell_i + \ell_{i-1} + a_{i-1} + 1 \geq b_{i-1} + 2 + \ell_i + \sqrt{n} \geq a_i + \ell_i + 1$$

and thus the rectangle R is sufficient.

We have an occurrence of L in A such that for every $i \in [q-1]$, the 1-entries of layer i are mapped to 1-entries in S_i. Thus all 1-entries of the q-th layer lie in $\mathcal{B}_q$. Since the widths of strips $\mathcal{B}_{q,j}$ are $\ell_q - 1$, the southwesternmost 1-entry α and the northeasternmost 1-entry ω of the q-th layer are mapped to different strips $\mathcal{B}_{q,j}$ and $\mathcal{B}_{q,j'}$. Then

$$r_\alpha - c_\alpha - (r_\omega - c_\omega) \geq 2(\sqrt{n} + \ell_q)$$
$$r_\alpha + c_\alpha + r_\omega + c_\omega \leq 2(b_{q-1} + \sqrt{n}).$$

and so

$$c_\alpha + r_\omega \leq b_{q-1} - \ell_q = a_{q-1} + \ell_{q-1} - 1,$$

which cannot occur since the ℓ_{q-1} 1-entries of the $(q-1)$-st layer are mapped to S_{q-1}.

□

1.4 Quadratic lower bound

Let

$$r_i := \begin{cases} i & \forall i \le k,\ i \text{ even} \\ 2k+1-i & \forall i \le k,\ i \text{ odd} \\ i & \forall i > k,\ i \text{ odd} \\ 2k+1-i & \forall i > k,\ i \text{ even.} \end{cases}$$

Given a positive integer k, let $\mathrm{Cross}(2k)$ be the $2k$-permutation matrix with 1-entries at positions (r_i, i) For example,

$$\mathrm{Cross}(8) = \begin{pmatrix} 0&0&0&0&0&0&0&1 \\ 0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0 \\ 1&0&0&0&0&0&0&0 \end{pmatrix}.$$

The construction of a matrix that avoids $\mathrm{Cross}(2k)$ is similar to the construction of matrices avoiding layered permutation matrices, but now, we iterate to depth.

The *distance between entries* $\alpha = (r, c)$ and $\beta = (r', c')$ is $\mathrm{dist}(\alpha, \beta) = |r - r'| + |c - c'|$. If $r - r'$ and $c - c'$ are nonzero and have opposite signs, we say that α and β are in the SW-NE direction. If the signs are nonzero and the same, α and β are in the NW-SE direction.

An *antistrip* $\bar{S}(a, b)$ in a matrix A is the set of elements (i, j) satisfying $a \le i + j \le b$. A *strip* $S(a, b)$ is the set of elements (i, j) satisfying $a \le i - j \le b$.

Let $A(1, n)$ be the $n \times n$ matrix with $2n-1$ 1-entries in the strip $S(0, 1)$. Then $A(1, n)$ avoids $\mathrm{Cross}(2)$.

For $k \ge 2$, we consider only such n that are the k^{th} power of an integer. Let $n' := n^{(k-1)/k}$, let $n'' := n^{(k-2)/k}$ and let $\Delta := n' + n'' + 4k$.

The matrix $A(k, n)$ has 1-entries in the strips $S^- := S(-(n' + 2k - 2), -n')$ and $S^+ := S(n', n' + 2k - 2)$. Let A' be the matrix $A(k-1, n')$ mirrored by the vertical axis. For every nonegative integer $i < \lfloor n/\Delta \rfloor$, we place a copy B_i of A' inside $A(k, n)$ to the intersection of rows $[i\Delta + 1, i\Delta + n']$ and columns $[i\Delta + 1, i\Delta + n']$. See Fig. 1.9.

Claim 1.23. *The matrix $A(k, n)$ avoids* $\mathrm{Cross}(2k)$.

Proof. We proceed by induction on k. The claim is trivial for $k = 1$. For a contradiction, assume that $A(k, n)$ contains $\mathrm{Cross}(2k)$ and fix one such occurrence. Let $\alpha_i = (r_i, i)$ be the 1-entry of $\mathrm{Cross}(2k)$ in column i and let $\phi(\alpha_i)$ be the 1-entry of $A(k, n)$ representing α_i. We can move $\phi(\alpha_1)$ to any 1-entry of $A(k, n)$ below and to the left of $\phi(\alpha_1)$. Thus, we can assume that $\phi(\alpha_1) \in S^+$ and, similarly, $\phi(\alpha_{2k}) \in S^-$.

Since α_2 and α_{2k-1} are at distance $2k - 1$ in the SW-NE direction from α_1, $\phi(\alpha_2)$ and $\phi(\alpha_{2k-1})$ are outside S^+. Because of their distance and the SW-NE direction from α_{2k}, both are outside S^- as well.

Let B_i and B_{i+j} be two of the copies of A', where $j > 0$. Every 1-entry of B_{i+j} is to the right and below every 1-entry of B_i. In addition, all 1-entries of $A(k-1, n')$ are inside the antistrip $\bar{S}(-n'' + 2k, n'' + 2k)$. We thus conclude that for every 1-entry β of B_i and β' of B_{i+j}, β and β' are the NW-SE direction and $\text{dist}(\beta, \beta') \geq 2\Delta - 2n'' - 4k = 2n' + 4k$.

We have $\text{dist}(\phi(\alpha_2), \phi(\alpha_{2k-1})) < \text{dist}(\phi(\alpha_1), \phi(\alpha_{2k}))$, because α_2 and α_{2k-1} lie inside the axis-parallel rectangle determined by α_1 and α_{2k}. Since $\phi(\alpha_1)$ and $\phi(\alpha_{2k})$ are in $\bar{S}(-n' - 2k + 2, n' + 2k - 2)$ and in the SW-NE direction, their distance is at most $2(n' + 2k)$.

Thus if $\phi(\alpha_2)$ and $\phi(\alpha_{2k-1})$ are not in the same B_i, then we reach a contradiction:

$$2(n' + 2k) \geq \text{dist}(\phi(\alpha_1), \phi(\alpha_{2k})) > \text{dist}(\phi(\alpha_2), \phi(\alpha_{2k-1}) \geq 2n' + 4k.$$

On the other hand, if $\phi(\alpha_2)$ and $\phi(\alpha_{2k-1})$ are in the same B_i, then all $\phi(\alpha_2)$, $\phi(\alpha_3)$, ..., $\phi(\alpha_{2k-1})$ are. We thus have an occurrence of Cross$(2k-2)$ in $A(k-1, n')$, a contradiction with the induction hypothesis. □

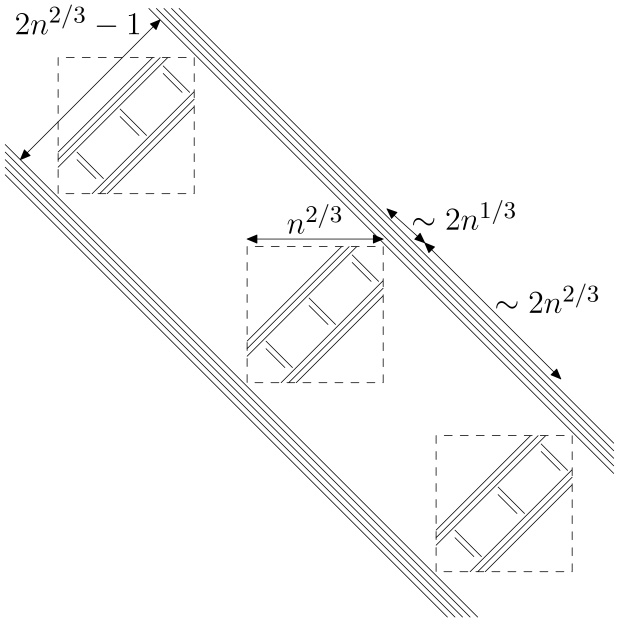


Figure 1.9: The matrix $A(3, n)$, which avoids Cross(6). Full lines represent diagonals with 1-entries.

Claim 1.24. *Let $k \geq 1$ and let n be a k-th power of an integer Then $A(k, n)$ contains at least $2k^2 n - O(n^{(k-1)/k})$ 1-entries.*

Proof. When $k = 1$, we have $h(k, n) \geq 2n - O(1)$, which was to be proven. Let $h(k, n)$ be the number of 1-entries in $A(k, n)$. The number of 1-entries in S^+ and S^- is $2(2k-1)(n - O(n'))$. The number of copies of $A(k-1, n')$ in $A(k, n)$ is

$\lfloor n/\Delta \rfloor$ and so

$$\begin{aligned} h(k,n) &\geq 2(2k-1)(n - O(n')) + h(k-1,n')\left(\frac{n}{\Delta} - 1\right) \\ &\geq 2(2k-1)n - O(n^{(k-1)/k}) \\ &\quad + \left(2(k-1)^2 n^{(k-1)/k} - O(n^{(k-2)/k})\right)\left(n^{1/k} - O(1)\right) \\ &\geq (4k-2)n + (2k^2 - 4k + 2)n - O(n^{(k-1)/k}) \\ &= 2k^2 n - O(n^{(k-1)/k}). \end{aligned}$$

□

Proof of Theorem 1.3. By Claim 1.23, for every k and n a k-th power of an integer, $A(k,n)$ avoids Cross$(2k)$ and by Claim 1.24, $A(k,n)$ has $2k^2 n - O(n^{(k-1)/k})$ 1-entries. Thus, for every even $k' \geq 2$, $c_{\text{Cross}(k')} \geq k'^2/2$. □

2. Permutations avoiding a given permutation

2.1 Introduction

Given a permutation matrix P, let $S_P(n)$ be the set of all n-permutation matrices avoiding P, $T_P(n)$ the set of all $n \times n$ $\{0,1\}$-matrices avoiding P and $T_P(n,m)$ the set of all $n \times n$ $\{0,1\}$-matrices containing exactly m 1-entries and avoiding P. Obviously

$$T_P(n) \supseteq T_P(n,n) \supseteq S_P(n).$$

The *Stanley–Wilf limit* of a permutation matrix P is defined as

$$s_P = \lim_{n\to\infty} \sqrt[n]{|S_P(n)|}.$$

The Stanley–Wilf conjecture was formulated by Stanley and Wilf around 1992 and asserts that s_P always exists and is finite. A weaker modification claimed that for any given P, $\sqrt[n]{|S_P(n)|}$ is bounded. Arratia [Arr99] showed that both versions are equivalent, that s_P always exists and that, when s_P is finite, then

$$\forall n \in \mathbb{N} : |S_P(n)| \leq (s_P)^n.$$

Let s_k be the maximum of s_P over all k-permutation matrices P.

Alon and Friedgut [AF00] used an upper bound on the length of generalized Davenport–Schinzel sequences to give a quasiexponential upper bound on $S_P(n)$ for every P.

Klazar [Kla00] showed that

$$|T_P(n)| \leq 15^{c_P n},$$

where c_P is the Füredi–Hajnal limit of P.

The Stanley–Wilf conjecture was then settled when Marcus and Tardos [MT04] proved the Füredi–Hajnal conjecture. The upper bound thus obtained was

$$s_k \leq 15^{2k^4\binom{k^2}{k}}.$$

By a result of Valtr published in [KK03], for every k and every k-permutation matrix P, $s_P \geq (k-1)^2/e^3$. There are infinitely many permutation matrices P with $s_P \geq 9.47(k-1)^2/9$ [Bón07].

Two permutations π and π' are *Wilf-equivalent* if $|S_\pi(n)| = |S_{\pi'}(n)|$ for every n. Two permutations whose permutation matrices become identical after transposing, mirroring or rotating one of them are trivially Wilf–equivalent. Given a t-permutation τ and an $(n-t)$-permutation π, the direct sum $\tau \oplus \pi$ of τ and π is the n-permutation $(\tau_1, \ldots, \tau_t, \pi_1 + t, \ldots, \pi_{n+t} + t)$. By a result of Backelin, West and Xin [BWX07], for every t and every permutation π, the permutations $(1, \ldots, t) \oplus \pi$ and $(t, \ldots, 1) \oplus \pi$ are Wilf-equivalent. See also a generalization by de Mier in Theorem 1.4. The Wilf-equivalence of the 4-permutations 4132 and 3142 was proved by Stankova [Sta96]. The list of known Wilf-equivalences is exhausted by the result of Stankova and West [SW02] stating that for every permutation π, $(2,3,1) \oplus \pi$ and $(3,1,2) \oplus \pi$ are Wilf-equivalent. Stankova and West [SW02] mention that no other pair of permutations of length at most seven is Wilf-equivalent. This is a consequence of computed values $|S_\pi(n)|$ for small n.

There are three classes of Wilf-equivalent 4-permutations. Their representants are 1234, 1342 and 1324. The exact formula for $|S_{1234}(n)|$ was found by Gessel [Ges90] and for $|S_{1342}(n)|$ by Bona [Bón97].

The third class resisted so far all attempts to enumerate it. Zeilberger believes [EV05] that "Not even God knows number of 1324-avoiders of length 1,000". The value of s_{1324} is not known either. In the recent years, the value of s_{1324} has been narrowed down to being at least 9.47 [AER$^+$06] and at most $7 + 4\sqrt{3} < 13.93$ [Bón12]. Madras and Liu [ML10] used Markov chain Monte Carlo methods to estimate the values $|S_{1324}(n+1)|/|S_{1324}(n)|$ for many different values of n in the range $[25, 155]$. Steingrímsson used a slight modification of this method to estimate $|S_{1324}(1001)|/|S_{1324}(1000)| \sim 11.01$. Extrapolations of the available numerical results [AER$^+$06, ML10, Ste12] suggest that the Stanley–Wilf limit is very likely to be somewhere between 11 and 12.

In Section 2.2, we improve the reduction from the Füredi–Hajnal conjecture to the Stanley–Wilf conjecture.

Theorem 2.1. *For every permutation matrix P*

$$s_P \leq 2.88 c_P^2 \quad \text{and thus for every } k \geq 1,$$
$$s_k \leq 2.88 \left(2k^4 \binom{k^2}{k}\right)^2.$$

In Section 2.3, we prove that the Füredi–Hajnal limit is bounded from above by a polynomial of the Stanley–Wilf limit. The proof relies on the ideas from the proof of the Füredi–Hajnal conjecture.

Theorem 2.2. *There is an absolute constant α such that for every permutation matrix P of size at least 2×2*

$$c_P \leq \alpha s_P^{4.5}.$$

In Section 2.4 we show that a large Füredi–Hajnal limit of some permutation matrix results in many permutation matrices avoiding another (larger) permutation matrix. The proof follows the approach used in Section 3.2.3.

Theorem 2.3. *For every $k \geq 2$ we have*

$$c_k \leq 8 s_{\lfloor 3k/2 \rfloor}.$$

These bounds mean that showing an upper bound polynomial in k on one of the constants c_k, s_k would give an upper bound polynomial in k on the other one.

The Stanley–Wilf limit is usually defined for permutations. We rephrased it in terms of permutation matrices, so the definitions satisfy $s_\pi = s_{P_\pi}$. To simplify the notation, we will sometimes use s_π instead of s_{P_π}.

Section 2.5 focuses on similar questions for higher-dimensional permutation matrices. A d-dimensional $\{0,1\}$-matrix P of size $k \times \cdots \times k$ is a *d-dimensional* *k-permutation matrix* if P contains k 1-entries and the positions of each two 1-entries of P differ in all coordinates. An extension of the Füredi–Hajnal conjecture to higher dimensions was proved by Klazar and Marcus [KM07]. For any given d-dimensional permutation matrix P, they showed that if a d-dimensional $n \times \cdots \times n$ $\{0,1\}$-matrix A avoids P, then A has at most $O(n^{d-1})$ 1-entries and there are such matrices A with $\Omega(n^{d-1})$ 1-entries.

The existence of an extension of the Stanley–Wilf conjecture to higher dimensions remains an open problem. For a d-dimensional permutation matrix P let $S_{P,d}(n)$ be the set of d-dimensional $n \times \cdots \times n$ permutation matrices avoiding P. In Section 2.5, we provide a small improvement over the trivial upper bound $|S_{P,d}(n)| \leq (n!)^{d-1}$.

Theorem 2.4. *For a fixed d-dimensional k-permutation matrix P*

$$n^{n(d-2+o(1))} \leq |S_{P,d}(n)| \leq n^{n\left(\frac{d(d-2)}{d-1}+o(1)\right)}.$$

The upper bound is obtained by a proof similar to the proof of Theorem 2.1.

2.2 Improved Klazar's reduction

Proof of Theorem 2.1. We can assume $c_P \geq 1$, since otherwise $s_P = 0$ and the statement is true.

The reduction is based on Klazar's reduction [Kla00]. We start with a 1×1 matrix $A_0 := (\bullet)$. In each step, we transform the matrix A_i of size $2^i \times 2^i$ into A_{i+1} of size $2^{i+1} \times 2^{i+1}$ by replacing each entry ω of A_i by a 2×2 block containing only 0-entries if and only if $\omega = 0$. There is a single possibility how to replace a 0-entry and fifteen possibilities of replacing a 1-entry. The number of 1-entries is non-decreasing, so we are only interested in matrices A_i with at most n 1-entries. Another estimate on the number of 1-entries uses the fact that if A_i contains P, then $A_{i+1}, A_{i+2} \dots$ contain P as well. So we consider only matrices A_i that avoid P, thus A_i has at most $\mathrm{ex}_P(2^i) \leq c_P \cdot 2^i$ 1-entries.

Phase 1: We use the estimate that the number of 1-entries in A_i is at most $c_P \cdot 2^i$ and get

$$|T_P(2^i)| \leq 15^{c_P \cdot 2^{i-1}} \cdot |T_P(2^{i-1})| \leq 15^{c_P \cdot (2^{i-1}+2^{i-2})} \cdot |T_P(2^{i-2})| \leq \cdots \leq 15^{c_P \cdot 2^i}. \quad (2.1)$$

Klazar continues until $2^i \geq n$, but we stop when $i = a$, which will be chosen later.

Phase 2: This time we use the estimate that the number of 1-entries in A_i is at most n. Using $a = \lfloor \log_2(n/c_P) \rfloor$, we could now easily show $s_P = O\left(c_P^{\log_2 15}\right)$, but our aim is a better estimate.

We will count how many transformations of matrices from $T_P(2^{a+i-1})$ give a matrix from $T_P(2^{a+i}, m)$. We define j_1, j_2, j_3, j_4 to be the numbers of 1-entries that were replaced by a block with 1, 2, 3, 4 1-entries, respectively. There are four possible replacements of a 1-entry that don't increase the number of 1-entries, six increase it by one, four by two and one by three. This gives the following recursive formula for the upper bound on $|T_P(2^{a+i}, m)|$:

$$\sum_{\substack{j_1,j_2,j_3,j_4\geq 0\\ j_1+2j_2+3j_3+4j_4=m}} \binom{m-j_2-2j_3-3j_4}{j_1,j_2,j_3,j_4} \cdot |T_P(2^{a+i-1}, m-j_2-2j_3-3j_4)| \cdot 4^{j_1}6^{j_2}4^{j_3}1^{j_4}$$

To simplify the computations, we define the function $u : \mathbb{N}_0 \times \mathbb{Z} \to \mathbb{N}_0$:

$$\begin{aligned}
u(0,m) &:= 1 \qquad \text{for every } m \geq 0\\
u(i,m) &:= 0 \qquad \text{when } m < 0\\
u(i,m) &:= \sum_{\substack{j_2,j_3,j_4\geq 0\\ j_2+j_3+j_4\leq m}} \binom{m}{m-j_2-j_3-j_4, j_2, j_3, j_4} \cdot 4^{m-2j_2-3j_3-4j_4}6^{j_2}4^{j_3}1^{j_4}\cdot\\
&\quad \cdot u(i-1, m-j_2-2j_3-3j_4). \text{ when } i>0 \text{ and } m \geq 0
\end{aligned}$$

We have $|T_P(2^{a+i}, m)| \leq u(i,m)|T_P(2^a)|$ because it is true for $i = 0$ and the differences between the recursive formulas are that the one for $u(i,m)$ adds several nonnegative summands and changes the multinomial coefficient. But, as one can check, the value of the multinomial coefficient never decreases.

For each nonnegative i, we will find some positive d_i such that for all integers m we will have $u(i,m) \leq (4^i d_i)^m$. First, $d_0 := 1$ satisfies the inequality for $i = 0$. For $i > 0$, if m is negative, the inequality is trivial, otherwise

$$\begin{aligned}
u(i,m) \leq &\sum_{\substack{j_2,j_3,j_4\geq 0\\ j_2+j_3+j_4\leq m}} \binom{m}{m-j_2-j_3-j_4, j_2, j_3, j_4}\cdot\\
&\cdot (4^{i-1}d_{i-1})^{m-j_2-2j_3-3j_4} \cdot 4^{m-2j_2-3j_3-4j_4} \cdot 6^{j_2}4^{j_3}1^{j_4}\\
=&(4^i d_{i-1})^m \sum_{\substack{j_1',j_2,j_3,j_4\geq 0\\ j_1'+j_2+j_3+j_4=m}} \binom{m}{j_1', j_2, j_3, j_4}\cdot\\
&\cdot \left(\frac{6}{d_{i-1}4^{i+1}}\right)^{j_2} \left(\frac{4}{d_{i-1}^2 4^{2i+1}}\right)^{j_3} \left(\frac{1}{d_{i-1}^3 4^{3i+1}}\right)^{j_4}\\
=&(4^i d_{i-1})^m \left(1 + \frac{6}{d_{i-1}4^{i+1}} + \frac{4}{d_{i-1}^2 4^{2i+1}} + \frac{1}{d_{i-1}^3 4^{3i+1}}\right)^m.
\end{aligned}$$

Thus we can set d_i to anything at least as large as

$$d_{i-1} \cdot \left(1 + 6/\left(d_{i-1}4^{i+1}\right) + 4/\left(d_{i-1}^2 4^{2i+1}\right) + 1/\left(d_{i-1}^3 4^{3i+1}\right)\right) \tag{2.2}$$

It follows that the sequence $d_0, d_1, \ldots,$ is nondecreasing. We will count d_1 and d_2 exactly and then estimate the rest.

For $i = 1$, the expression (2.2) becomes 1.44140625, so we can set $d_1 := 1.4415$.

Then we get $d_2 \geq 1.537989\ldots$, so we let $d_2 := 1.538$. For $i \geq 3$ let

$$\begin{aligned} d_i &= d_{i-1} \cdot \left(1 + \frac{6}{d_{i-1}4^{i+1}} + \frac{4}{d_{i-1}^2 4^{2i+1}} + \frac{1}{d_{i-1}^3 4^{3i+1}}\right) \\ &\leq d_{i-1} \cdot \left(1 + \frac{6}{d_2 4^{i+1}} + \frac{4}{d_2^2 4^{2i+1}} + \frac{1}{d_2^3 4^{3i+1}}\right) \\ &\leq d_{i-1} \exp\left(\frac{6}{d_2 4^{i+1}} + \frac{4}{d_2^2 4^{2i+1}} + \frac{1}{d_2^3 4^{3i+1}}\right) \\ &\leq d_2 \prod_{j=3}^{i} \left(\exp\left(\frac{6}{d_2 4^{j+1}} + \frac{4}{d_2^2 4^{2j+1}} + \frac{1}{d_2^3 4^{3j+1}}\right)\right) \\ &= d_2 \exp\left(\sum_{j=3}^{i} \frac{6}{d_2 4^{j+1}} + \sum_{j=3}^{i} \frac{4}{d_2^2 4^{2j+1}} + \sum_{j=3}^{i} \frac{1}{d_2^3 4^{3j+1}}\right) \\ &\leq d_2 \exp\left(\frac{4}{3}\frac{6}{d_2 4^4} + \frac{16}{15}\frac{4}{d_2^2 4^7} + \frac{64}{63}\frac{1}{d_2^3 4^{10}}\right) \\ &\leq 1.57. \end{aligned}$$

Let $d_\infty := 1.57$. All in all, we have just proven that for any i and m:

$$|T_P(2^{a+i}, m)| \leq 4^{im} d_\infty^m |T_P(2^a)| \leq 4^{im} d_\infty^m \cdot 15^{c_P \cdot 2^a}, \tag{2.3}$$

where the last inequality follows from (2.1). We could finish when $2^{a+i} \geq n$ for the first time, which would already result in $s_P = O(c_P^2)$, but to achieve a better multiplication constant, we continue until $a + i$ equals some b such that $2^b \geq 2n^2$.

Every n-permutation matrix avoiding P can be expanded by adding empty rows and columns to form a matrix from $T_P(2^b, n)$. This can be done in $\binom{2^b}{n}^2$ ways while the reverse process is unique — we just delete all empty rows and columns and take what remains. Therefore $|T_P(2^b, n)| \geq |S_P(n)|\binom{2^b}{n}^2$.

Since $2^b \geq 2n^2$, we can estimate:

$$\binom{2^b}{n} \geq \frac{(2^b - n)^n}{n!} \geq \frac{2^{b\cdot n}\left(1 - \frac{1}{2n}\right)^n}{en\left(\frac{n}{e}\right)^n} \geq \frac{2^{b\cdot n} \cdot e^{-1}}{en\left(\frac{n}{e}\right)^n}.$$

We now have

$$\begin{aligned} |S_P(n)| &\leq |T_P(2^b, n)| \cdot \binom{2^b}{n}^{-2} \\ &\leq 4^{n\cdot(b-a)} \cdot d_\infty^n \cdot 15^{c_P \cdot 2^a} \cdot \left(en\left(\frac{n}{e}\right)^n \cdot e \cdot 2^{-b\cdot n}\right)^2 \\ &= e^4 n^2 \left(4^{b-a} d_\infty \frac{n^2}{e^2} 4^{-b}\right)^n 15^{c_P \cdot 2^a} \end{aligned}$$

and so

$$\begin{aligned} \sqrt[n]{|S_P(n)|} &\leq \sqrt[n]{e^4 n^2} \frac{d_\infty}{e^2} n^2 4^{-a} 15^{c_P 2^a / n} \\ &= \sqrt[n]{e^4 n^2} \frac{d_\infty}{e^2} 4^{-a} \exp\left(2\ln(n) + \frac{\ln(15) c_P 2^a}{n}\right). \end{aligned}$$

Let $g_a(n) := 2\ln(n) + \ln(15)c_P 2^a/n$. A routine calculation shows that for any given $a > 0$, $g_a(n)$ has its minimum at $n = \ln(15)c_P 2^{a-1}$ and is decreasing on the interval $(0, \ln(15)c_P 2^{a-1})$. So we will set

$$n(a) := \left\lfloor \ln(15)c_P 2^{a-1} \right\rfloor \tag{2.4}$$

and estimate

$$g_a(n(a)) \leq g_a\left(\ln(15)c_P 2^{a-1} - 1\right) \leq g_a\left(\ln(15)c_P 2^{a-1}\left(1 - \frac{1}{2^a}\right)\right).$$

Since $\lim_{a\to\infty} n(a) = \infty$ and from [Arr99] $\lim_{n\to\infty} \sqrt[n]{|S_P(n)|}$ exists, we obtain

$$\begin{aligned}
s_P &= \lim_{n\to\infty} \sqrt[n]{|S_P(n)|} = \lim_{a\to\infty} \sqrt[n(a)]{|S_P(n(a))|} \\
&\leq \lim_{a\to\infty}\left(\sqrt[n(a)]{e^4 n(a)^2}\right)\frac{d_\infty}{e^2}\lim_{a\to\infty}\left(4^{-a}\exp\left(g_a(n(a))\right)\right) \\
&\leq 1\cdot\frac{d_\infty}{e^2}\lim_{a\to\infty}4^{-a}\exp\left(g_a\left(\ln(15)c_P 2^{a-1}\left(1-\frac{1}{2^a}\right)\right)\right) \\
&= \frac{d_\infty}{e^2}\lim_{a\to\infty}4^{-a}\exp\left(2\ln\left(\ln(15)c_P 2^{a-1}\left(1-\frac{1}{2^a}\right)\right) + \frac{\ln(15)c_P 2^a}{\ln(15)c_P 2^{a-1}\left(1-\frac{1}{2^a}\right)}\right) \\
&\leq \frac{d_\infty}{e^2}\lim_{a\to\infty}4^{-a}\left(\ln(15)c_P 2^{a-1}\right)^2\left(1-\frac{1}{2^a}\right)^2\exp\left(\frac{2}{1-\frac{1}{2^a}}\right) \\
&= \frac{d_\infty}{e^2}\cdot\frac{\ln^2(15)}{4}{c_P}^2\lim_{a\to\infty}4^{-a}4^a\left(1-\frac{1}{2^a}\right)^2\exp\left(\frac{2}{1-\frac{1}{2^a}}\right) \\
&= d_\infty\cdot\frac{\ln^2(15)}{4}{c_P}^2 \\
&\leq 2.88{c_P}^2.
\end{aligned}$$

Theorem 1 from [Arr99] now gives

$$\forall n \geq 1 : |S_P(n)| \leq (2.88c_P^2)^n.$$

□

Remark. We also have $\sqrt[n]{|T_P(n,n)|} \leq O({c_P}^2)$. This follows from Equation (2.3) with $m = n(a)$ and $i = \lceil \log(n(a)) - a \rceil$ using $n(a)$ as defined in (2.4). However, $\lim_{n\to\infty} \sqrt[n]{|T_P(n)|} \geq 2^{c_P}$. To show this, we take an $n \times n$ $\{0,1\}$-matrix A with $\mathrm{ex}_P(n)$ 1-entries that avoids P. The matrix A contains $2^{\mathrm{ex}_P(n)}$ different $n \times n$ matrices and all these matrices avoid P.

2.3 Bounding the extremal function using the Stanley–Wilf limit

Lemma 2.5. *Let P be any permutation matrix and let B be a matrix of size $b \times c$ containing at least b 1-entries in each row. If B avoids P, then*

$$|S_P(b)| \geq \left(\frac{b^2}{e^2 c}\right)^b.$$

Proof. We take the rows of B one by one from top to bottom and from each of them, we select some 1-entry in a column that was not used previously. This way, we constructed a b-permutation matrix contained in B, thus avoiding P. This construction gives us at least $b!$ occurrences of b-permutation matrices, but some can be different occurrences of the same matrix. To count the largest possible number of occurrences of a given b-permutation matrix, we observe that the rows are given but we can select any b-tuple out of the c columns. All in all, the number of different b-permutation matrices avoiding P is at least

$$\frac{b!}{\binom{c}{b}} \geq \frac{\left(\frac{b}{e}\right)^b}{\left(\frac{ce}{b}\right)^b} = \left(\frac{b^2}{e^2 c}\right)^b.$$

□

Lemma 2.6. *For a given permutation matrix P take any $l \in \mathbb{N}$ such that $\sqrt[7]{l}$ is an integer and*

$$|S_P(l^{10/7})| < \left(\frac{l^{6/7}}{2e^2}\right)^{l^{10/7}}.$$

Then

$$\forall n \in \mathbb{N} : \mathrm{ex}_P(n) \leq \left(2l^{27/7} + 10l^{24/7} + 8l^2\right) n.$$

Proof. First observe that if P has size 1×1, then the lemma holds.

By a theorem of Arratia [Arr99], if P has size at least 2×2, then for every $i, j \geq 1$ we have $|S_P(i+j)| \geq |S_P(i)| \cdot |S_P(j)|$. Extending this, we have $|S_p(\alpha \cdot i)| \geq |S_p(i)|^\alpha$, and so the conditions of the lemma also imply

$$|S_P(l)| < \left(\frac{l^{6/7}}{e^2}\right)^l \qquad \text{and} \qquad |S_P(l^{8/7})| < \left(\frac{l^{6/7}}{e^2}\right)^{l^{8/7}}.$$

Let $A = (a_{i,j})$ be any $n \times n$ permutation matrix avoiding P. We start similarly to the proof of the Füredi–Hajnal conjecture [MT04] — we cut the matrix A by horizontal and vertical cuts into a grid of blocks $K_{i,j}$ of sizes $2l^2 \times 2l^2$ and discard the incomplete blocks on the right and at the bottom. That is,

$$K_{i,j} := \left\{a_{i',j'} : i' \in \{2l^2 i + 1, \ldots, 2l^2(i+1)\},\ j' \in \{2l^2 j + 1, \ldots, 2l^2(j+1)\}\right\}.$$

The j^{th} *column of blocks* is $\mathcal{C}_j := \{K_{i,j} : i \in \{0, 1, \ldots, \lfloor n/(2l^2) \rfloor - 1\}\}$ and the i^{th} *row of blocks* is $\mathcal{R}_i := \{K_{i,j} : j \in \{0, 1, \ldots, \lfloor n/(2l^2) \rfloor - 1\}\}$. We say that a block is *wide* if it contains more than l nonzero columns, *very wide* if it contains more than $l_1 = l^{8/7}$ nonzero columns and *ultrawide* if it contains more than $l_2 = l^{10/7}$ nonzero columns. Similarly, a block is *tall, very tall, ultratall* if it has more than l, l_1, l_2 nonzero rows, respectively. Throughout the proof we will use the following observation:

Claim 2.7. *We take b blocks from the same column of blocks and separately contract the columns of each of them. This way we obtain a $b \times 2l^2$ matrix $B = (b_{i,j})$ with one row for each block, such that $b_{i,j} = 0$ if and only if the i^{th} selected block contains no 1-entry in its j^{th} column. If B contains P, then A contains P as well.*

Proof. For each 1-entry in the occurrence of P in the contracted matrix B, we take any 1-entry from the column from which it was contracted. Because P is a permutation matrix, the relative positions of these 1-entries do not change and they form an occurrence of P in the original matrix A. □

We now finish the proof of Lemma 2.6. If $n \leq 2l^{27/7}$, the claim is trivial, otherwise we count the maximal number of 1-entries in a matrix A that avoids P:

- The discarded blocks have together at most $2 \cdot 2l^2 n$ 1-entries.

- Each nonzero block which is neither wide nor tall, has at most l^2 1-entries. As was shown in [MT04], if we contract each block of A into a single element (whose value is 1 exactly if the block is nonzero), we obtain a matrix that avoids P. So the number of nonzero blocks is at most $\mathrm{ex}_P(\lfloor n/2l^2 \rfloor)$ and this value can be estimated from the induction hypothesis.

- Each ultrawide or ultratall block has at most $4l^4$ 1-entries. We will show that there are fewer than $l_2 = l^{10/7}$ ultrawide blocks in one column of blocks and fewer than l_2 ultratall blocks in one row of blocks. It is enough to prove this only for ultrawide blocks; the proof for ultratall blocks is the same. For contradiction, suppose there are at least l_2 ultrawide blocks in the same column of blocks. We contract the columns of each of them as in Claim 2.7 and obtain a $l_2 \times 2l^2$ matrix B with l_2 rows each of which has at least l_2 1-entries. Lemma 2.5 then gives

$$\left|S_P\left(l^{10/7}\right)\right| \geq \left(\frac{l_2^2}{2e^2 l^2}\right)^{l_2} = \left(\frac{l^{6/7}}{2e^2}\right)^{l^{10/7}},$$

which contradicts the conditions of Lemma 2.6.

- Each very wide or very tall block which is neither ultrawide nor ultratall has at most $l_2^2 = l^{20/7}$ 1-entries. To count the maximal number of very wide blocks in one column of blocks we first contract each such block to a row with at least $l_1 = l^{8/7}$ 1-entries. If some l_1 consecutive rows have all their 1-entries in at most l_2 columns of A, we will remove all the other columns and obtain an $l_1 \times l_2$ matrix with at least l_1 1-entries in each row and consequently

$$\left|S_P\left(l^{8/7}\right)\right| \geq \left(\frac{l_1^2}{e^2 l_2}\right)^{l_1} = \left(\frac{l^{6/7}}{e^2}\right)^{l^{8/7}},$$

which is not possible and so there are at least l_2 nonzero columns in each group of l_1 consecutive rows. Contracting this group gives a row with at least l_2 1-entries and as was previously shown, there are fewer than l_2 such rows. We conclude that there are fewer than $l_2 l_1 = l^{18/7}$ very wide blocks in one column of blocks.

- In each wide or tall block which is neither very wide nor very tall, there are at most $l_1^2 = l^{16/7}$ 1-entries. We divide the wide blocks into groups of

l consecutive blocks. If all the 1-entries in the blocks of one group lied in only l_1 columns, there would be at least

$$|S_P(l)| \geq \left(\frac{l^2}{e^2 l_1}\right)^l = \left(\frac{l^{6/7}}{e^2}\right)^l$$

l-permutation matrices avoiding P. So each group can be contracted into a row with at least l_1 1-entries. But as we have shown, there are fewer than $l_2 l_1$ such rows and therefore there are at most $l_2 l_1 l = l^{25/7}$ wide blocks in one column of blocks.

The overall number of 1-entries is at most

$$\begin{aligned} \mathrm{ex}_P(n) &\leq 2 \cdot 2l^2 n + l^2 \,\mathrm{ex}_P\left(\left\lfloor \frac{n}{2l^2} \right\rfloor\right) + 2\left(4l^4 l^{10/7} + l^{20/7} l^{18/7} + l^{16/7} l^{25/7}\right) \frac{n}{2l^2} \\ &\leq \left(4l^2 + l^{27/7} + 5l^{24/7} + 4l^2 + 4l^{24/7} + l^{24/7} + l^{27/7}\right) n \\ &\leq \left(2l^{27/7} + 10l^{24/7} + 8l^2\right) n. \end{aligned}$$

□

Proof of Theorem 2.2. We take the smallest $l > (2e^2 s_P)^{7/6}$ that is a seventh power of an integer. Because $(2e^2 s_P)^{7/6} \geq 1$, we will find a suitable l not larger than $2^7 (2e^2 s_P)^{7/6}$. For every integer i, the number of i-permutation matrices avoiding P is at most s_P^i and from the choice of l, $s_P^i < \left(\frac{l^{6/7}}{2e^2}\right)^i$. Thus we can use Lemma 2.6 and substituting $l \leq 2^7 (2e^2 s_P)^{7/6}$ into its result finishes the proof:

$$c_P \leq \left(2^{32.5} e^9 s_P^{4.5} + 5 \cdot 2^{29} e^8 s_P^4 + 2^{58/3} e^{14/3} s_P^{7/3}\right).$$

□

2.4 Sets of permutations from matrices

Let P be a permutation matrix. A pair of consecutive rows $(i, i+1)$ is an *increasing pair* if the 1-entry in the $(i+1)$-st row has larger column index than the 1-entry in the i-th row. Otherwise, $(i, i+1)$ is a *decreasing pair*. These definitions correspond to an ascent and a descent in the inverse of the permutation corresponding to P. An *increasing (decreasing) chain of rows* is a sequence of consecutive rows such that each consecutive pair of rows is increasing (decreasing).

Let r be the number of increasing pairs in P. We transform P into a $(k+r)$-permutation matrix P^+, the *interleaved* P. For every increasing pair $(i, i+1)$, we add a zero row in between the rows i and $i+1$. Subsequently, we add r zero columns after the columns of P. The matrix P^+ is obtained by placing the $r \times r$ identity matrix (or any other permutation matrix) on the intersection of the new rows and columns. The 1-entries, rows and columns of P^+ present already in P are the *original 1-entries, rows and columns.* Then for every increasing pair $(i, i+1)$ of rows of P^+, the row $i+1$ is one of the newly added rows. Thus no two original rows are in the same increasing chain of rows.

A *word matrix* is a matrix with exactly one 1-entry in every column. Given an $m \times n$ word matrix H, we order its 1-entries primarily by increasing row index and secondarily by increasing column index. The *expansion* X of H is the n-permutation matrix with 1-entries on positions (i, j) such that H has its ith 1-entry in column j. This can be viewed as inflating every row to form the identity matrix. See also Fig. 3.4.

Claim 2.8. *If a word matrix H avoids P then the expansion X of H avoids P^+.*

Proof. For every i, we say that the ith 1-entry of H *corresponds* to the ith 1-entry of X.

For a contradiction, assume that X contains P^+. Let U be the set of the k 1-entries in the occurrence of P^+ in X that correspond to the original 1-entries of P^+. Let U' be the set of 1-entries of H corresponding to the 1-entries in U. The 1-entries of U form an occurrence of P in X. We show that the 1-entries of U' form an occurrence of P in H.

The 1-entries in both X and H are ordered primarily by their row. Therefore if we prove that the rows of the 1-entries of U' are all distinct, then they form an occurrence of P in H.

By the definition of the inflation, the 1-entries of H lying in the same row correspond to 1-entries on a chain of increasing rows of X. Since no two original rows of P^+ are in the same increasing chain of rows of P^+, the rows of the 1-entries of U' are all distinct. □

Lemma 2.9. *Let P be a k-permutation matrix with r increasing pairs of consecutive rows. Let n be an integer. If there is an $n \times n$ P-avoiding matrix A with c 1-entries in every column, then*

$$s_{k+r} \geq c/4.$$

Proof. Let P^+ be the interleaved P. Let $\mathcal{H}$ be the set of at least c^n $n \times n$ word matrices contained in A. Each matrix $H \in \mathcal{H}$ avoids P.

By Claim 2.8, the expansion X of every $H \in \mathcal{H}$ avoids P^+, the interleaved P. Let $\mathcal{X}$ be the set of expansions of matrices $H \in \mathcal{H}$. For every $X \in \mathcal{X}$, there are at most 2^n word matrices H' with n columns and no zero row and whose expansion is X. Indeed, each row of H' corresponds to a nonempty interval of consecutive rows of X and there are 2^{n-1} ways to partition the rows of X into nonempty intervals. There are at most 2^n ways to turn each H' into an $n \times n$ word matrix by adding zero rows. Thus there are at most 4^n $n \times n$ word matrices H whose expansion is X and so

$$|\mathcal{X}| \geq \frac{|\mathcal{H}|}{4^n} \geq \left(\frac{c}{4}\right)^n.$$

Since P^+ is a $(k + r)$-permutation matrix and every $X \in \mathcal{X}$ avoids P^+,

$$s_{k+r} \geq s_{P^+} \geq c/4.$$

□

Proof of Theorem 2.3. Let $\bar{P}$ be a k-permutation matrix such that $c_{\bar{P}} = c_k$.

For every $\varepsilon > 0$ we take a large enough integer n and an $n \times n$ matrix $\bar{A}$ that avoids $\bar{P}$ and has $(c_k - \varepsilon)n$ 1-entries. Let $c' = c_k - \varepsilon$.

A *light row (column)* is a row (column) with fewer than $c'/2$ 1-entries. We create a square matrix B from $\bar{A}$ by the following procedure. While the matrix has at least one light row and at least one light column, then we remove one light row and one light column. The process stops before all the rows and columns are removed, since it is impossible to remove $c'n$ 1-entries in n steps each of which removes fewer than $2c'/2$ 1-entries.

Thus we have an $n' \times n'$ matrix A' such that it either has no light row or no light column. If B has no light column, then we let P' be the transpose of $\bar{P}$ and A' be the transpose of B. Otherwise, let $P' := \bar{P}$ and $A' := B$.

If P' has fewer increasing pairs than decreasing pairs, then let $P := P'$ and $A := A'$. Otherwise, we mirror P' by the vertical axis to get P and A' to get A. This changes every decreasing pair in P' to an increasing pair in P and vice versa. Therefore P has at most $\lfloor k/2 \rfloor$ increasing pairs.

The matrix A is a P-avoiding $n' \times n'$ matrix with at least $c'/2$ 1-entries in every column. Thus we may use Lemma 2.9 with $c = c'/2 = (c_k - \varepsilon)/2$ and $r < \lfloor k/2 \rfloor$ and so, for every $\varepsilon > 0$,

$$s_{\lfloor 3k/2 \rfloor} \geq s_{k+r} \geq s_{P^+} \geq c/4 = (c_k - \varepsilon)/8.$$

□

Remark. In Section 1.4, we constructed an $n \times n$ matrix A that avoids a $2k$-permutation matrix Cross($2k$) and has $2k^2$ 1-entries in all but $o(n)$ columns. Let $k' = 3k$. A small modification of Lemma 2.9 shows that the interleaved k'-permutation matrix Cross$(2k)^+$ satisfies $s_{\text{Cross}(2k)^+} \geq k'^2/18$. This is better than the general lower bound $s_P \geq k'^2/e^3$ by Valtr published in the paper of Kaiser and Klazar [KK03]. However, since Cross$(2k)^+$ contains the $k \times k$ identity matrix, we have $s_{\text{Cross}(2k)^+} \geq (k'-1)^2/9$.

2.5 Higher-dimensional matrices

We call $M \in \{0,1\}^{[n_1]\times\cdots\times[n_d]}$ a *d-dimensional $\{0,1\}$-matrix of size $n_1 \times \cdots \times n_d$*.

A d-dimensional $\{0,1\}$-matrix P of size $k \times \cdots \times k$ is a *d-dimensional k-permutation matrix* if P contains k 1-entries and the positions of each two 1-entries of P differ in all coordinates.

We say that a d-dimensional $\{0,1\}$-matrix $P = (p_{i_1,\ldots,i_d})$ of size $k_1 \times \cdots \times k_d$ is *contained* in a d-dimensional $\{0,1\}$-matrix $A = (a_{i_1,\ldots,i_d})$ of size $n_1 \times \cdots \times n_d$ if there exist d increasing injections $f_i : [k_i] \to [n_i]$, $i = 1, 2, \ldots, d$ such that for all $i_1, i_2, \ldots, i_d \in [k] : p_{i_1,\ldots,i_d} = 1$ implies $a_{f_1(i_1),\ldots,f_d(i_d)} = 1$. If P is not contained in A, we say that A *avoids* P.

For a d-dimensional k-permutation matrix P and $a, b \in [d]$, let the *(a,b)-projection* of P, $\text{proj}_{a,b}(P)$, be the (2-dimensional) k-permutation matrix P' with $p'_{i,j} = 1$ exactly if P has a 1-entry whose a^{th} coordinate has value i and b^{th} coordinate has value j.

Klazar and Marcus [KM07] proved that for a fixed d-dimensional k-permutation matrix P, the maximum number of 1-entries in a d-dimensional matrix A of size $n \times \cdots \times n$ that avoids P is $\text{ex}_{P,d}(n) = \Theta(n^{d-1})$. This generalizes the Füredi–Hajnal conjecture.

Let P be a given d-dimensional k-permutation matrix P. Define $S_{P,d}(n)$ to be the set of all d-dimensional n-permutation matrices avoiding P and $T_{P,d}(n,m)$

to be the set of all d-dimensional matrices of size $n \times \cdots \times n$ that avoid P and have at most m 1-entries. Obviously $T_{P,d}(n,n) \supseteq S_{P,d}(n)$.

Proof of Theorem 2.4. The lower bound follows from the following observation. Let P and A be permutation matrices and let $P' := \mathrm{proj}_{1,2}(P)$ and $A' := \mathrm{proj}_{1,2}(A)$. If A' avoids P', then A avoids P.

For the given matrix P, we take a matrix A' that avoids P'. For any 2-dimensional n-permutation matrix A', there are $(n!)^{d-2} = n^{n(d-2+o(1))}$ d-dimensional n-permutation matrices A such that $A' = \mathrm{proj}_{1,2}(A)$. Because A' avoids P', all such matrices A are in $S_{P,d}(n)$.

The proof of the upper bound is similar to the proof of Theorem 2.1. We start with A_0, the $1 \times \cdots \times 1$ matrix containing one 1-entry. In each step, we transform the matrix A_i of size $2^i \times \cdots \times 2^i$ into A_{i+1} of size $2^{i+1} \times \cdots \times 2^{i+1}$ by replacing each 0-entry of A_i by a $2 \times \cdots \times 2$ block containing only 0-entries and each 1-entry of A_i by a $2 \times \cdots \times 2$ block containing at least one 1-entry. There is a single possibility how to replace a 0-entry and $2^{2^d} - 1$ possibilities of replacing a 1-entry. However only 2^d of the possible replacements of the 1-entry do not increase the number of 1-entries.

In the first phase we use the above mentioned estimate $\mathrm{ex}_{P,d}(2^i) = \Theta(2^{i(d-1)})$ from [KM07]. Thus $\mathrm{ex}_{P,d}(2^i) \le c_{P,d} 2^{i(d-1)}$ for some constant $c_{P,d}$ and

$$|T_{P,d}(2^i, n)| \le 2^{2^d \cdot c_{P,d} \cdot 2^{(i-1)(d-1)}} \cdot |T_{P,d}(2^{i-1}, n)| \le \cdots \le 2^{2^d \cdot c_{P,d} \cdot 2^{i(d-1)}}.$$

We stop when $i = a$, where $a := \lceil \log_2(n^{1/(d-1)}) \rceil$. Then $|T_{P,d}(2^a, n)| \le 2^{O(n)}$.

In the second phase which consists of $b := \lceil \log_2(n) \rceil - a \le \log_2(n)(d-2)/(d-1)+1$ steps, we use the fact that all matrices A_i have at most n 1-entries. During this phase, we will do at most bn replacements of a 1-entry, but only at most n of them will increase the number of 1-entries.

For each matrix A_{a+b} that was created from A_a, we order the replacements of 1-entries during the second phase primarily by the step in which the replacement occurred and secondarily by the lexicographic order of the position of the 1-entry being replaced. The $2^{2^d} - 1$ types of replacements of a 1-entry are assigned numbers from $[2^{2^d} - 1]$ so that the types of replacements that do not increase the number of 1-entries get numbers from $[2^d]$. Each matrix A_{a+b} that was created from A_a, is assigned a vector whose elements are from $[2^{2^d} - 1]$. The entries of the vector represent the replacements of 1-entries in the above defined order, the value of each entry is the type of the replacement. The vector has length at most bn and at most n its entries are from $[2^{2^d} - 1] \setminus [2^d]$. Then we append several entries to the vector to obtain a vector of length $(b+1)n$ with exactly n entries from $[2^{2^d} - 1] \setminus [d]$. Different matrices A_{a+b} created from the same matrix A_a get different vectors and thus

$$\begin{aligned} |T_{P,d}(2^{a+b}, n)| &\le 2^{2^d n} 2^{dbn} \binom{(b+1)n}{n} |T_{P,d}(2^a, n)| \\ &\le 2^{O(n)} 2^{\log_2(n) \frac{d(d-2)}{d-1} n} ((\log_2(n) + 2)e)^n \\ &\le n^{n\left(\frac{d(d-2)}{d-1} + o(1)\right)}. \end{aligned}$$

We have $|T_{P,d}(n,n)| \leq |T_{P,d}(2^{a+b}, n)|$, because $2^{a+b} \geq n$ and thus all matrices from $T_{P,d}(n,n)$ are in a one-to-one correspondence with matrices from $T_{P,d}(2^{a+b}, n)$ that have 0-entries in all places with some coordinate larger than n. □

Remark. There are more higher-dimensional generalizations of permutations. This section considered only the *sparse permutation matrices* in the terminology used, for example, by Eriksson and Linusson [EL00].

A d-dimensional matrix of size $k \times \cdots \times k$ is a *dense d-dimensional k-permutation matrix* if every 1-dimensional subarray of length k contains exactly one 1-entry. Thus there are k^{d-1} 1-entries altogether. For example, dense 3-dimensional permutation matrices correspond to Latin squares. It is not known what is the number of dense d-dimensional permutation matrices, only a $((1 + o(1))n/e^{d-1})^{n^{d-1}}$ upper bound is known [LL11] and a lower bound of the same type for $d = 3$ [vLW92].

Forbidding dense permutation matrices is much less restrictive than forbidding sparse ones. For example, already for $d = 3$ and every $k \geq 2$, there are $n \times n \times n$ matrices with $n^{3-1/k}$ 1-entries avoiding all 3-dimensional k-permutation matrices. Indeed, if a 3-dimensional matrix M contains a dense 3-dimensional k-permutation, then $\mathrm{proj}_{1,2}(M)$ contains the $k \times k$ matrix $K_{k,k}$ with all entries equal to 1. We can thus create M by taking n copies of an $n \times n$ $K_{k,k}$-avoiding matrix with $n^{2-1/k}$ 1-entries and putting them on top of each other.

It is not known how many dense d-dimensional permutation matrices avoid a fixed d-dimensional permutation matrix for $d \geq 3$. An occurrence of a $k \times k \times k$ permutation matrix in an $n \times n \times n$ permutation matrix corresponds to a $k \times k$ Latin subsquare in the corresponding $n \times n$ Latin square. Computational evidence suggest, that for every n and $k \geq 3$, a majority of $n \times n$ Latin squares has no $k \times k$ Latin subsquare [MW99].

2.6 Conclusions and open problems

We have shown that the values s_P and c_P are closely related. Another, rather weak, evidence of this connection is the following computation result of Madras and Liu [ML10]: They used Markov chain random process to construct random 1324-avoiding permutations. The plot of one of these permutations of length 100 [ML10, Figure 9] shows some similarities with the construction of a 1324-avoiding matrix in Section 1.3.

It is impossible to find an increasing function $g : \mathbb{R} \to \mathbb{R}$ such that $s_P = g(c_P)$. This can be seen on the example of the permutation matrices I and F corresponding to the permutations 1234 and 1342, respectively. We have $c_I = c_F = 6$ (see Chapter 1), but $s_I = 9$ and $s_F = 8$ (see Section 2.1). The following is still plausible:

Problem 1. *Does there exist a constant r_1, such that for all permutation matrices P*

$$s_P \geq r_1 c_P^2 \text{ ?}$$

Several authors conjectured that the Stanley–Wilf limit is maximized by some layered pattern. Some even conjectured which layered patterns attain the largest

Stanley–Wilf limit and those were the ones with short layer lengths. See for example the survey of Steingrímsson [Ste12].

However, the layered k-permutation matrix with the largest extremal function is the one with layers of lengths 1, $k-2$ and 1. Moreover, when k is even, the matrix Cross(k) has even larger extremal function. The author believes that this holds for the Stanley–Wilf limits as well.

Conjecture 1. *For every $k \geq 4$, the layered permutations with layer lengths 1, $k-2$ and 1 is the easiest to avoid of all layered k-permutations. That is, the Stanley–Wilf limit of $\rho(1, k-2, 1)$ is larger than the Stanley–Wilf limit of every other layered k-permutation.*

Conjecture 2. *There is a k_0 such that for every even $k \geq k_0$, the Stanley–Wilf limit of* Cross(k) *is larger than the Stanley–Wilf limit of every layered k-permutation.*

Results computed by the author using the method of Madras and Liu [ML10] with Steingrímsson's [Ste12] modification provide a small evidence in favor of Conjecture 1[1].

Section 2.5 contains an improvement of the trivial upper bound on $|S_{P,d}(n)|$, but the gap between the bounds still remains very large. An extension of the Stanley–Wilf conjecture could be the following.

Conjecture 3. *Given any d-dimensional permutation matrix P, there exists a constant $s_{P,d}$ such that*

$$|S_{P,d}(n)| \leq s_{P,d}^n \cdot n^{(d-2)n}.$$

[1] The author estimated $|S_\pi(101)|/|S_\pi(100)|$ for each of 153426, 154326 and 132546. Four independent runs of length 10^7 and four of length 10^8 were done and the average number of ways to add a new last element to the permutations of the runs to make a π-avoiding permutation of length 101 was counted. In the case $\pi = 153426$, the average in each of the runs was between 24.49 and 24.58 and between 24.52 and 24.55 for the runs of length 10^8. In the case $\pi = 154326$, the average was always between 26.81 and 27 and between 26.89 and 26.92 for the runs of length 10^8. In the case $\pi = 132546$, the average was always between 24.17 and 24.33 and between 24.27 and 24.29 for the runs of length 10^8.

3. Sets of permutations with bounded VC-dimension

3.1 Introduction

Let $\mathcal{T}$ be a set system on $[n] = \{1, 2, \ldots, n\}$. We say that a set $K \subset [n]$ is *shattered* by $\mathcal{T}$ if every subset of K appears as an intersection of K and some set from $\mathcal{T}$. The *Vapnik–Chervonenkis dimension (VC-dimension)* of $\mathcal{T}$ is the size of the largest set shattered by $\mathcal{T}$. Sauer's lemma gives the exact value of the maximum size of a set system on $[n]$ with VC-dimension k, which is a polynomial in n of degree k. More on the VC-dimension of sets systems and its history can be found for example in [Mat02].

Motivated by the so-called acyclic linear orders problem, Raz [Raz00] defined the VC-dimension of a set $\mathcal{P}$ of permutations: Let S_n be the set of all *n-permutations*, that is, permutations of $[n]$. The *restriction of $\pi \in S_n$ to the k-tuple $(a_1, a_2, \ldots, a_k)$ of positions* (where $1 \le a_1 < a_2 < \cdots < a_k \le n$) is the k-permutation π' satisfying $\forall i, j : \pi'(i) < \pi'(j) \Leftrightarrow \pi(a_i) < \pi(a_j)$. The k-tuple of positions $(a_1, \ldots, a_k)$ is *shattered by* $\mathcal{P}$ if each k-permutation appears as a restriction of some $\pi \in \mathcal{P}$ to $(a_1, \ldots, a_k)$. The *VC-dimension of* $\mathcal{P}$ is the size of the largest set of positions shattered by $\mathcal{P}$. Let $r_k(n)$ be the size of the largest set of n-permutations with VC-dimension k.

Raz [Raz00] proved that $r_2(n) \le C^n$ for some constant C and asked whether an exponential upper bound on $r_k(n)$ can also be found for every $k \ge 3$.

An n-permutation π avoids a k-permutation ρ if none of the restrictions of π to a k-tuple of positions is ρ. So the set of permutations avoiding $\rho \in S_k$ has VC-dimension smaller than k. Thus, Raz's question generalizes the Stanley–Wilf conjecture discussed in Chapter 2.

We show in Section 3.2 that the size of a set of n-permutations with VC-dimension k cannot be much larger than exponential in n. The result has an application in enumerating the number of weak isomorphism classes of simple complete topological graphs [Kyn12]. Let $\alpha(n)$ be the inverse of the Ackermann function; see Section 3.2.2 for its definition.

Theorem 3.1. *The sizes of sets of permutations with bounded VC-dimension satisfy*

$$\begin{aligned} r_3(n) &\le \alpha(n)^{(4+o(1))n}, \\ r_4(n) &\le 2^{n\cdot(2\alpha(n)+3\log_2(\alpha(n))+O(1))}, \\ r_{2t+2}(n) &\le 2^{n\cdot((2/t!)\alpha(n)^t+O(\alpha(n)^{t-1}))} && \text{for every } t \ge 2 \text{ and} \\ r_{2t+3}(n) &\le 2^{n\cdot((2/t!)\alpha(n)^t\log_2(\alpha(n))+O(\alpha(n)^t)} && \text{for every } t \ge 1. \end{aligned}$$

On the other hand, we give a negative answer to Raz's question in Section 3.3.

Theorem 3.2. *We have*

$$r_3(n) \geq (\alpha(n)/2 - O(1))^n \qquad \textit{and}$$
$$r_{2t+3}(n) \geq r_{2t+2}(n) \geq 2^{n\cdot((1/t!)\alpha(n)^t - O(\alpha(n)^{t-1}))} \qquad \textit{for every } t \geq 1.$$

An *n-permutation matrix* is an $n \times n$ $\{0,1\}$-matrix with exactly one 1-entry in every row and every column. Permutations and permutation matrices are in a one-to-one correspondence that assigns to a permutation π a permutation matrix A_π with $A_\pi(i,j) = 1 \Leftrightarrow \pi(j) = i$.

We consider the following modification of the question of Füredi and Hajnal considered in Chapter 1. We study the maximal number $p_k(n)$ of 1-entries in an $n \times n$ matrix such that no $(k+1)$-tuple of columns contains all $(k+1)$-permutation matrices. It can be easily shown that $p_2(n) = 4n - 4$. Indeed, consider an $n \times n$ matrix with at least $4n - 3$ 1-entries. Remove the highest and the lowest 1-entry in every column. Then the first and the last row of the resulting matrix contain no 1-entry and thus one of the rows contains three 1-entries. The three columns of the original matrix containing these 1-entries contain every 3-permutation matrix. The lower bound $4n - 4$ can be achieved for example by filling the two top rows and some two columns with 1's.

Theorem 3.3. *We have*

$$2n\alpha(n) - O(n) \leq p_3(n) \leq O(n\alpha(n)),$$
$$p_{2t+2}(n) = n2^{(1/t!)\alpha(n)^t \pm O(\alpha(n)^{t-1})} \quad \textit{for every } t \geq 1 \textit{ and}$$
$$n2^{(1/t!)\alpha(n)^t - O(\alpha(n)^{t-1})} \leq p_{2t+3}(n) \leq n2^{(1/t!)\alpha(n)^t \log_2(\alpha(n)) + O(\alpha(n)^t)} \quad \textit{for every } t \geq 1.$$

The upper bounds from Theorem 3.3 are proven as Corollary 3.8 in Section 3.2.1 and the lower bounds as Corollary 3.26 in Section 3.3.2.

Let $\lambda_s(n)$ be the maximum length of a Davenport–Schinzel sequence over n symbols. The following are the currently best known bounds on $\lambda_s(n)$.

$$2n\alpha(n) - O(n) \leq \lambda_3(n) \leq 2n\alpha(n) + O\left(n\sqrt{\alpha(n)}\right),$$
$$n \cdot 2^{(1/t!)\alpha(n)^t - O\left(\alpha(n)^{t-1}\right)} \leq \lambda_{2t+2}(n) \leq n \cdot 2^{(1/t!)\alpha(n)^t + O\left(\alpha(n)^{t-1}\right)} \qquad \text{for } t \geq 1 \text{ and}$$
$$n \cdot 2^{(1/t!)\alpha(n)^t - O\left(\alpha(n)^{t-1}\right)} \leq \lambda_{2t+3}(n) \leq n \cdot 2^{(1/t!)\alpha(n)^t \log_2 \alpha(n) + O\left(\alpha(n)^t\right)} \qquad \text{for } t \geq 1.$$

The upper bound on λ_3 is by Klazar [Kla99], the lower bounds on λ_s for $s > 3$ are by Agarwal, Sharir and Shor [ASS89] and all the other bounds were proved by Nivasch [Niv10].

Pettie [Pet12] recently announced the following improved bounds:

$$\Omega(n\alpha(n)2^{\alpha(n)}) \leq \lambda_5(n) \leq O(n\alpha^2(n)2^{\alpha(n)}) \qquad \text{and}$$
$$\lambda_{2t+3}(n) \leq n \cdot 2^{(1/t!)\alpha(n)^t(1+o(1))} \qquad \text{for } t \geq 2.$$

Our proofs are based on several results on Davenport–Schinzel sequences as well as on sequences with other forbidden patterns. The results on sequences that we use are mentioned in more detail in Sections 3.2.1, 3.2.2 and 3.3.1, where they are transformed into claims about matrices with forbidden patterns.

An *s-partition of the rows* of an $m \times n$ matrix M is a partition of the interval of integers $\{1, \ldots, m\}$ into s intervals $\{1 = m_1, \ldots, m_2 - 1\}$, $\{m_2, \ldots, m_3 - 1\}$, $\ldots$, $\{m_s, \ldots, m = m_{s+1} - 1\}$. A matrix M contains a *B-fat* (r, s)*-formation* if there exists an s-partition of the rows and an r-tuple of columns each of which has B 1-entries in each interval of rows. Note that the order of the columns in the matrix is not important for this notion. See Fig. 3.1 for an example of a 1-fat $(3, 4)$-formation. In Section 3.2.2, we prove the following lemma, which gives an upper bound on the number of 1-entries an $n \times n$ matrix can have and still not contain any B-fat (B, s)-formation. It is used in the proof of Theorem 3.1 in Section 3.2.3, analogously to the use of Raz's Technical Lemma [Raz00].

Lemma 3.4. *For all positive integers s, n and B, an $n \times n$ matrix M with at least $\zeta_s(n)Bn$ 1-entries contains a B-fat (B, s)-formation, where $\zeta_s(n)$ are functions of the form*

$$\begin{aligned}
\zeta_2(n) &= O(1),\\
\zeta_3(n) &= O(\alpha(n)),\\
\zeta_4(n) &= O(\alpha(n)^2),\\
\zeta_5(n) &= O(\alpha(n)2^{\alpha(n)}),\\
\zeta_{2t+3}(n) &= 2^{(1/t!)\alpha(n)^t + O(\alpha(n)^{t-1})} \qquad \text{for } t \geq 2 \text{ and}\\
\zeta_{2t+4}(n) &= 2^{(1/t!)\alpha(n)^t \log(\alpha(n)) + O(\alpha(n)^t)} \qquad \text{for } t \geq 1.
\end{aligned}$$

More generally, for all positive integers m, n, s and B, an $m \times n$ matrix M with at least $\zeta_s(m)Bn$ 1-entries contains a B-fat $(\lfloor nB/m \rfloor, s)$-formation.

The proof of the lemma is based on a proof of the upper bound on the number of symbols in the so-called formation-free sequences (see definition in Section 3.2.1) from Nivasch's paper [Niv10].

By an argument similar to the proof of $p_2(n) \leq 4n - 4$ above, it is easy to verify that every $m \times n$ matrix M with at least $3n$ 1-entries contains a 1-fat $(\lceil n/m \rceil, 3)$-formation. A similar result for 2-fat formations would slightly improve the upper bounds on $r_3(n)$ and $r_4(n)$.

Problem 2. *Does there exist a constant c such that for every m and n, every $m \times n$ matrix M with at least cn 1-entries contains a 2-fat $(\lfloor n/m \rfloor, 3)$-formation?*

All logarithms in this chapter are in base 2.

3.2 Upper bounds

3.2.1 Numbers of 1-entries in matrices

A sequence S of length l over an alphabet Γ is a function $S : [l] \to \Gamma$. An (r, s)*-formation* is a sequence formed by s concatenated permutations of the same r-tuple of symbols. The permutations in a formation are its *troops*. A sequence $S = (a_1, \ldots, a_l)$ is r*-sparse* if $a_i \neq a_j$ whenever $0 < |i - j| < r$.

An (r, s)*-formation-free sequence* is a sequence that is r-sparse and contains no (r, s)-formation as a subsequence. Let $F_{r,s}(n)$ be the maximum length of an (r, s)-formation-free sequence over n symbols. Formation-free sequences were first studied by Klazar [Kla92].

To be able to use results on sequences for matrices, we use the *matrix→sequence transcription* MST (our name) defined by Pettie [Pet11a] who improved an earlier transcription by Füredi and Hajnal [FH92]. The letters of the sequence correspond to the columns of the matrix. The matrix is transcribed row by row from top to bottom. Let Seq_{i-1} be the sequence created from the first $i-1$ rows. We consider the set C_i of letters corresponding to the columns having a 1-entry in the row i. The letters in C_i are ordered in the order of the last appearance in Seq_{i-1}; the one that appeared last in Seq_{i-1} is first and so on. The letters that did not appear in Seq_{i-1} are ordered arbitrarily and placed after those that did appear. The ordered sequence C_i is then appended to Seq_{i-1}. The length of the resulting sequence $\mathrm{MST}(M) = \mathrm{Seq}_m$ is equal to the number of 1-entries in M and the size of the alphabet is n. Note that the previous papers ([FH92, Pet11a]) transcribe the matrices column by column instead of row by row.

A *block* in a sequence is a contiguous subsequence containing only distinct symbols. Note that $\mathrm{MST}(M)$ can be decomposed into m or fewer blocks.

A set S of rs 1-entries forms an (r, s)*-formation* in M if there exists an s-partition of the rows and an r-tuple of columns each of which has a 1-entry of S in every interval of rows of the partition. See Fig. 3.1. In this and all other figures, circles and full circles represent the 1-entries and empty space represents the 0-entries. A matrix M is (r, s)*-formation-free* if it contains no (r, s)-formation.

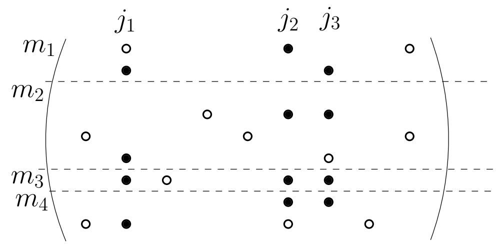


Figure 3.1: A $(3, 4)$-formation on columns j_1, j_2 and j_3. Full circles represent the 1-entries of the formation. Empty circles represent 1-entries outside of this formation.

Lemma 3.5. *A* $\{0, 1\}$*-matrix* M *contains an* (r, s)*-formation if and only if the sequence* $\mathrm{MST}(M)$ *contains an* (r, s)*-formation.*

Proof. Observe that an (r, s)-formation in a matrix M implies an (r, s)-formation in $\mathrm{MST}(M)$.

The proof of the other direction is more complicated, because symbols of one block of $\mathrm{MST}(M)$ may be present in two troops of the (r, s)-formation in $\mathrm{MST}(M)$. To overcome this complication, we consider such an (r, s)-formation in $\mathrm{MST}(M)$, whose each troop ends earliest possible. Assume that the i-th troop ends with an occurrence of a symbol a in the j-th block of $\mathrm{MST}(M)$ and that the $(i + 1)$-st troop begins with b from the j-th block. Since a precedes b in the j-th block, we know, by the definition of $\mathrm{MST}(M)$, that a appears somewhere between the occurrences of b and a of the i-th troop. Therefore, the i-th troop could end earlier, contradicting the selection of the (r, s)-formation. $\square$

Nivasch gives the following upper bound on the maximum length $F_{r,s}(n)$ of an (r, s)-formation-free sequence on n symbols:

Theorem 3.6. ([Niv10, Theorem 1.3]) *For every $r \in \mathbb{N}$*

$$F_{r,4}(n) \leq O(n\alpha(n)).$$

For every r and every $s \geq 5$, letting $t := \lfloor (s-3)/2 \rfloor$, we have

$$F_{r,s}(n) \leq n2^{(1/t!)\alpha(n)^t \log(\alpha(n))+O(\alpha(n)^t)} \qquad \text{when } s \text{ is even and}$$
$$F_{r,s}(n) \leq n2^{(1/t!)\alpha(n)^t+O(\alpha(n)^{t-1})} \qquad \text{when } s \text{ is odd.}$$

Let $p'_k(n)$ be the maximum number of 1-entries in an $(k+1, k+1)$-formation-free $n \times n$ matrix. Theorem 3.6 implies the following upper bounds on $p'_k(n)$.

Lemma 3.7. *We have*

$$p'_3(n) \leq O(n\alpha(n)).$$

For every fixed $k \geq 4$, letting $t := \lfloor (k-2)/2 \rfloor$, we have

$$p'_k(n) \leq n2^{(1/t!)\alpha(n)^t \log(\alpha(n))+O(\alpha(n)^t)} \qquad \text{when } k \text{ is odd and}$$
$$p'_k(n) \leq n2^{(1/t!)\alpha(n)^t+O(\alpha(n)^{t-1})} \qquad \text{when } k \text{ is even.}$$

Proof. Take a $(k+1, k+1)$-formation-free matrix M. Then $\mathrm{MST}(M)$ does not contain any $(k+1, k+1)$-formation by Lemma 3.5.

The sequence $\mathrm{MST}(M) = a_1, a_2, \ldots, a_p$ can be made $(k+1)$-sparse by removing at most kn occurrences of symbols. Indeed, whenever two occurrences a_i, a_j (where $i < j$) of the same symbol appear at distance at most k, then a_i is among the last k symbols preceding the block containing a_j. Thus, it suffices to take the blocks from left to right and in each of them remove the at most k symbols that appear as the last k symbols preceding the block. The resulting sequence is thus a $(k+1, k+1)$-formation-free sequence of length differing by $O(n)$ from the number of 1-entries of M. The result then follows from Theorem 3.6. $\square$

This proves the upper bounds in Theorem 3.3 by observing that a $(k+1)$-tuple of columns with a $(k+1, k+1)$-formation contains every $(k+1)$-permutation matrix.

Corollary 3.8. *For every fixed $k \geq 3$ if we let $t := \lfloor (k-2)/2 \rfloor$, then*

$$p_3(n) \leq O(n\alpha(n)),$$
$$p_k(n) \leq n2^{(1/t!)\alpha(n)^t \log(\alpha(n))+O(\alpha(n)^t)} \qquad \text{when } k \text{ is odd and greater than 3 and}$$
$$p_k(n) \leq n2^{(1/t!)\alpha(n)^t+O(\alpha(n)^{t-1})} \qquad \text{when } k \text{ is even.}$$

3.2.2 Fat formations in matrices

A sequence S is an $\mathrm{AFF}_{r,s,k}(m)$*-sequence*[1] if it contains no (r,s)-formation as a subsequence, can be decomposed into m or fewer blocks and each symbol of the sequence appears at least k times. Let $\Pi'_{r,s,k}(m)$ be the maximum number of symbols in an $\mathrm{AFF}_{r,s,k}(m)$-sequence.

Let $\alpha_d(m)$ be the dth function in the *inverse Ackermann hierarchy*. That is, $\alpha_1(m) = \lceil m/2 \rceil$, $\alpha_d(1) = 0$ for $d \geq 2$ and $\alpha_d(m) = 1 + \alpha_d(\alpha_{d-1}(m))$ for $m, d \geq 2$. The *inverse Ackermann function* is defined as $\alpha(m) := \min\{k : \alpha_k(m) \leq 3\}$.

Nivasch defines a hierarchy of functions $R_s(d)$, which we shift by 1 in the index. That is, our $R_s(d)$ is the original $R_{s-1}(d)$. We thus have the functions defined for $s \geq 2$ and $d \geq 2$. The values are $R_2(d) = 2$, $R_3(d) = 3$, $R_4(d) = 2d+1$, $R_s(2) = 2^{s-2} + 1$ and

$$R_s(d) = 2(R_{s-1}(d)-1)+(R_{s-2}(d)-1)(R_s(d-1)-3)+1 \qquad \text{when } s \geq 5 \text{ and } d \geq 3.$$

For $s \geq 5$, if we let $t = \lfloor (s-3)/2 \rfloor$, then $R_s(d) = 2^{(1/t!)d^t \log(d) + O(d^t)}$ if s is even and $R_s(d) = 2^{(1/t!)d^t + O(d^{t-1})}$ when s is odd.

Lemma 3.9. ([Niv10, Corollary 5.14]) *For every $d \geq 2$, $s \geq 3$, $r \geq 2$, m and k satisfying $m \geq k \geq R_s(d)$ we have*

$$\Pi'_{r,s,k}(m) \leq c'_s r m \alpha_d(m)^{s-3},$$

where c'_s is a constant depending only on s.

The linear dependence of the upper bound on r is not explicitly mentioned in [Niv10], but can be revealed from the proof. In the base case, the dependence on r is linear (Lemmas 5.9 and 5.10 in [Niv10]) and in Recurrences 5.11 and 5.13, the right-hand side can be rewritten as r times an expression not depending on r.

It was shown [KV94, Pet11b] that doubling letters in the forbidden subsequence usually has small impact on the maximum length of a generalized DS-sequence. Geneson [Gen09] generalized the linear upper bound from the Füredi–Hajnal conjecture to forbidden double permutation matrices. We show a similar behavior of formation-free sequences and matrices. For $s \geq 2$, a set S of $r(2s-2)$ 1-entries forms a *doubled* (r,s)*-formation* in M if there exists an s-partition of the rows and an r-tuple of columns each of which has one 1-entry of S in the top and bottom interval of rows of the partition and two 1-entries in every other interval. A matrix M is *doubled* (r,s)*-formation-free* if it contains no doubled (r,s)-formation. A $\mathrm{DFF}_{r,s,k}(m)$*-matrix* is a doubled (r,s)-formation-free matrix with m rows and at least k 1-entries in every column. Let $\Delta_{r,s,k}(m)$ be the maximum number of columns in a $\mathrm{DFF}_{r,s,k}(m)$-matrix.

In Corollary 3.16 we show an analogue of Lemma 3.9 for doubled (r,s)-formation-free matrices. The proof follows the structure of the proof of Corollary 5.14 in [Niv10]. First, we show some simple bounds on $\Delta_{r,s,k}(m)$. The case $d = 2$ of Corollary 3.16 is proved in Corollary 3.14 by Recurrence 3.13 and the remaining cases follow from Recurrence 3.15. Corollary 3.16 will give a sequence of upper bounds on $\Delta_{r,s,k}(m)$. Typically, the bounds are superlinear in m for r,

[1] AFF is an abbreviation for *almost-formation-free*.

s and k fixed and the subsequence of bounds applicable is limited by the values of s and k. As k grows (keeping r and s fixed) the best applicable bound gets closer and closer to linear. When one lets k be a suitable function of $\alpha(m)$, the bound becomes linear in m.

If $m < k$, no matrix with m rows can have k 1's in every column.

Observation 3.10. *For every r, s, k, m, if $m < k$, then*

$$\Delta_{r,s,k}(m) = 0.$$

Observation 3.11. *For every r, s, k, m, if $k < 2s - 2$, then*

$$\Delta_{r,s,k}(m) = \infty.$$

Analogously to [Niv10, Lemma 5.10], all the other values of $\Delta_{r,s,k}(m)$ are finite.

Observation 3.12. *For every $r \geq 2$, $s \geq 2$ and $m \geq 2s - 2$*

$$\Delta_{r,s,2s-2}(m) \leq (r-1)\binom{m-s+1}{s-1} \leq rm^{s-1}.$$

Proof. If each column in an r-tuple of columns has the same position of the 2nd, 4th, ..., $(2s-2)$nd 1-entry, then the first $2s-2$ 1-entries from the columns form a doubled (r, s)-formation. □

Recurrence 3.13. *For every r, k, m and $s \geq 3$*

$$\Delta_{r,s,2k+1}(2m) \leq 2\Delta_{r,s,2k+1}(m) + 2\Delta_{r,s-1,k}(m).$$

Proof. As in the proof of [Niv10, Recurrence 5.11], we consider a $\mathrm{DFF}_{r,s,2k+1}(2m)$-matrix M and cut its rows into the upper m rows and the lower m rows. The *local columns* are those with all 1-entries in the same half of rows. There are at most $2\Delta_{r,s,2k+1}(m)$ local columns. Columns that are not local are *global*. Consider the submatrix M_1' formed by the upper half of rows of global columns with at least half of their 1's in the upper half of rows. Let M_1 be the matrix created from M_1' by removing the lowest 1 in every column of M_1'. If M_1 contains a doubled $(r, s-1)$-formation, then M contains a doubled (r, s)-formation. Thus M_1 has at most $\Delta_{r,s-1,k}(m)$ columns. A symmetric argument can be applied on the global columns with at least half of their 1's in the lower half of rows. □

Corollary 3.14. *For every fixed $s \geq 2$ and for all integers r, k, m satisfying $k \geq 2^{s-1} + 2^{s-2} - 1$ we have*

$$\Delta_{r,s,k}(m) \leq \bar{c}_s rm \log(m)^{s-2},$$

where $\bar{c}_s$ is a constant depending only on s.

Proof. The proof proceeds by induction on s and m. The base case of $s = 2$ follows from Observation 3.12 and the cases with $m < k$ from Observation 3.10. Recurrence 3.13 is used as the induction step. □

Recurrence 3.15. *For every nonnegative r, m, k_1, k_2, k_3, k_4 and t satisfying $m > t$, $k_1 \geq k_2+1 \geq 2$ and $k_4 \geq k_3 \geq 3$, if we let $k = 2k_1+(k_2+1)(k_3-3)+(k_4-k_3)+1$, then*

$$\Delta_{r,s,k}(m) \leq \left(1+\frac{m}{t}\right)(\Delta_{r,s,k}(t)+2\Delta_{r,s-1,k_1}(t)+\Delta_{r,s-2,k_2}(t))+ \\ +\Pi'_{r,s,k_3}\left(1+\frac{m}{t}\right)+\Delta_{r,s,k_4}\left(1+\frac{m}{t}\right) \qquad \textit{for } s \geq 4 \textit{ and}$$

$$\Delta_{r,s,k}(m) \leq \left(1+\frac{m}{t}\right)(\Delta_{r,s,k}(t)+2\Delta_{r,s-1,k_1}(t)+r-1)+ \\ +\Pi'_{r,s,k_3}\left(1+\frac{m}{t}\right)+\Delta_{r,s,k_4}\left(1+\frac{m}{t}\right) \qquad \textit{for } s = 3.$$

Proof. Consider a $\mathrm{DFF}_{r,s,k}(m)$-matrix M. We partition the rows of M into $b := \lceil m/t \rceil \leq m/t+1$ layers $L_1, \ldots, L_b$ of at most t consecutive rows each.

A column is

- *local* in layer L_i if all its 1's appear in layer L_i,
- *top-concentrated* in layer L_i if it has at least k_1+1 1's in layer L_i and at least one 1-entry below L_i,
- *bottom-concentrated* in layer L_i if it has at least k_1+1 1's in layer L_i and at least one 1-entry above L_i,
- *middle-concentrated* in layer L_i if it has at least k_2+2 1's in layer L_i and at least one 1-entry above and one below layer L_i,
- *doubly-scattered* if it has at least two 1's in at least k_3 layers,
- *scattered* if it has a 1-entry in at least k_4 layers.

These categories are analogous to those used by Nivasch [Niv10], except that we added the category of doubly-scattered columns. This allows us to use Π' instead of Δ in one summand of the recurrence. As one of the consequences, when $s \geq 6$, the upper bound on the maximum number of 1's in a doubled (r,s)-formation-free $n \times n$ matrix in Lemma 3.4 is similar to the best known upper bound on $F_{r,s}(n)$, although it is closer to $F_{r,s+1}(n)$ when $s = 3$.

Every column falls into one of these categories. If a column is in none of them, then its number of 1's is maximized when it has k_1 1's in its top and bottom nonzero layers, k_2+1 1's in some other k_3-3 layers and a single 1 in some additional k_4-k_3 layers. Thus it contains only at most $2k_1+(k_2+1)(k_3-3)+(k_4-k_3) \leq k-1$ 1-entries.

For each layer L_i, the number of columns local in L_i is at most $\Delta_{r,s,k}(t)$. For every fixed i we consider the columns that are top-concentrated in L_i and let M'_i be the submatrix of M defined by these columns and the rows of L_i. Let M_i be obtained from M'_i by removing the lowest 1-entry from every column. If M_i contains a doubled $(r,s-1)$-formation, then M contains a doubled (r,s)-formation. Thus there are at most $\Delta_{r,s-1,k_1}(t)$ columns top-concentrated in L_i. Similarly, there are at most $\Delta_{r,s-1,k_1}(t)$ columns bottom-concentrated in L_i. For $s \geq 4$, there at most $\Delta_{r,s-2,k_2}(t)$ columns middle-concentrated in L_i. For $s = 3$, there are at most $r-1$ columns middle concentrated in L_i, because an r-tuple of columns with two 1's in layer L_i and at least one 1 above and one below contains a doubled $(r, 3)$-formation.

To bound the number of doubly-scattered columns, we *contract* each layer into a single row. That is, we write 1 for every column containing at least two 1's in the layer and 0 otherwise. If there is an (r, s)-formation on the contracted doubly-scattered columns, then M contains a doubled (r, s)-formation. Thus, by Lemma 3.5, there are at most $\Pi'_{r,s,k_3}(\lceil m/t \rceil)$ doubly-scattered columns. By a similar argument, the number of scattered columns is at most $\Delta_{r,s,k_4}(\lceil m/t \rceil)$. The only difference is that while contracting, we write 1 for the columns containing at least one 1 in the layer. □

Similarly to Nivasch's functions $R_s(d)$, we define a hierarchy of functions $D_s(d)$, where $s \geq 1$ and $d \geq 2$, as follows: $D_1(d) = 0$, $D_2(d) = 2$, $D_s(2) = 2^{s-1} + 2^{s-2} - 1$ and when $s, d \geq 3$

$$D_s(d) = 2D_{s-1}(d) + (D_{s-2}(d) + 1)(R_s(d-1) - 3) + D_s(d-1) - R_s(d-1) + 1.$$

Then

$$\begin{aligned} D_3(d) &= 2d + 1, \\ D_4(d) &\leq O(d^2), \\ D_5(d) &\leq O(d2^d), \\ D_{2t+3}(d) &\leq 2^{(1/t!)d^t + O(d^{t-1})} \quad \text{for } t \geq 2 \text{ and} \\ D_{2t+4}(d) &\leq 2^{(1/t!)d^t \log(d) + O(d^t)} \quad \text{for } t \geq 1. \end{aligned}$$

Corollary 3.16. *For every $d \geq 2$, $s \geq 2$, $r \geq 2$, m and k satisfying $m \geq k \geq D_s(d)$ we have*

$$\Delta_{r,s,k}(m) \leq c_s r m \alpha_d(m)^{s-2},$$

where c_s is a constant depending only on s.

Proof. The proof proceeds by induction on d, s and m similarly to the proof of [Niv10, Corollary 4.12]. In the case $s = 2$, we apply Observation 3.12 and so the lemma holds with $c_2 = 1$. For every $s \geq 3$ let $m_0(s)$ be a constant such that

$$m \geq 1 + (6s)^s \lceil \log_2(m) \rceil^{s^2} \qquad \text{for every } m \geq m_0(s).$$

Let $\widehat{c}_1 = \widehat{c}_2 = 1$ and for $s \geq 3$ we define $\widehat{c}_s$ in the order of increasing s as

$$\widehat{c}_s := \max\{c'_s, \bar{c}_s, 9\widehat{c}_{s-1}, 9\widehat{c}_{s-2}, m_0(s)^{s-1}\},$$

where $\bar{c}_s$ is the constant from Corollary 3.14 and c'_s is the constant from Lemma 3.9. For every $s \geq 3$ and $d \geq 2$, we define a function $\bar{\alpha}_{d,s}$ by $\bar{\alpha}_{2,s}(m) = \lceil \log(m) \rceil$, $\bar{\alpha}_{d,s}(m) = 1$ if $m \leq m_0(s)$ and

$$\bar{\alpha}_{d,s}(m) = 1 + \bar{\alpha}_{d,s}(6s\bar{\alpha}_{d-1,s}(m)^{s-2}) \qquad \text{otherwise.}$$

Then $\bar{\alpha}_{d,s}(m)$ is well defined and differs by at most an additive constant (depending on s) from the values of the dth inverse Ackermann function $\alpha_d(m)$ for all

s, d and m (this can be shown similarly to [Niv10, Appendix C]). The functions also satisfy $\bar{\alpha}_{d,s}(m) \geq \bar{\alpha}_{d,s-1}(m)$. It is thus enough to prove

$$\Delta_{r,s,k}(m) \leq \widehat{c}_s r m \bar{\alpha}_{d,s}(m)^{s-2}.$$

The case $d = 2$ follows from Corollary 3.14. The cases $m \leq m_0(s)$ follow from Observation 3.12. Now $s \geq 3$, $d \geq 3$ and $m > m_0(s)$. We apply Recurrence 3.15 with:

$$\begin{aligned} &k_1 = D_{s-1}(d), \qquad k_2 = D_{s-2}(d), \qquad k_3 = R_s(d-1), \\ &k_4 = D_s(d-1), \qquad k = D_s(d) \qquad \text{and} \qquad t = 6s\bar{\alpha}_{d-1,s}(m)^{s-2}. \end{aligned}$$

By the induction hypothesis,

$$\begin{aligned} 2\Delta_{r,s-1,k_1}(t) + \Delta_{r,s-2,k_2}(t) &\leq r\frac{\widehat{c}_s}{3} t\bar{\alpha}_{d,s}(m)^{s-3} \qquad \text{when } s \geq 4, \\ 2\Delta_{r,s-1,k_1}(t) + r - 1 &\leq r\frac{\widehat{c}_s}{3} t\bar{\alpha}_{d,s}(m)^{s-3} \qquad \text{when } s = 3, \\ \Delta_{r,s,k_4}\left(1 + \frac{m}{t}\right) &\leq \widehat{c}_s r \frac{2m}{t} \bar{\alpha}_{d-1,s}(m)^{s-2} \\ &\leq r\frac{\widehat{c}_s}{3s} \leq r\frac{\widehat{c}_s}{9} m\bar{\alpha}_{d,s}(m)^{s-3} \text{ for } s \geq 3 \end{aligned}$$

and by Lemma 3.9,

$$\Pi'_{r,s,k_3}\left(1 + \frac{m}{t}\right) \leq rs\widehat{c}_s \frac{2m}{t} \bar{\alpha}_{d-1,s}(m)^{s-3} \leq r\frac{\widehat{c}_s}{3} m \leq r\frac{\widehat{c}_s}{3} m\bar{\alpha}_{d,s}(m)^{s-3}.$$

Substituting into Recurrence 3.15 we get

$$\begin{aligned} \Delta_{r,s,k}(m) &\leq \frac{m}{t}\Delta_{r,s,k}(t) + \Delta_{r,s,k}(t) + (m+t) r\frac{\widehat{c}_s}{3}\bar{\alpha}_{d,s}(m)^{s-3} + \frac{4\widehat{c}_s}{9} rm\bar{\alpha}_{d,s}(m)^{s-3} \\ &\leq \frac{m}{t}\Delta_{r,s,k}(t) + \frac{7\widehat{c}_s}{9} rm\bar{\alpha}_{d,s}(m)^{s-3} + \Delta_{r,s,k}(t) + \frac{\widehat{c}_s}{3} rt\bar{\alpha}_{d,s}(m)^{s-3}. \end{aligned}$$

By Observation 3.12, $\Delta_{r,s,k}(t) \leq rt^{s-1} \leq r(6s\bar{\alpha}_{d-1,s}(m)^{s-2})^{s-1}$, which is at most rm, because $m \geq m_0(s)$. So $\Delta_{r,s,k}(t) \leq \widehat{c}_s rm/9$. Similarly $t\bar{\alpha}_{d,s}(m)^{s-3} \leq m/3$. Thus

$$\begin{aligned} \Delta_{r,s,k}(m) &\leq \frac{m}{t}\Delta_{r,s,k}(t) + \frac{7\widehat{c}_s}{9} rm\bar{\alpha}_{d,s}(m)^{s-3} + \frac{\widehat{c}_s}{9} rm + \frac{\widehat{c}_s}{9} rm \\ &\leq \frac{m}{t}\Delta_{r,s,k}(t) + \widehat{c}_s rm\bar{\alpha}_{d,s}(m)^{s-3} \\ &\leq \frac{m}{t}\widehat{c}_s rt\bar{\alpha}_{d,s}(t)^{s-2} + \widehat{c}_s rm\bar{\alpha}_{d,s}(m)^{s-3} \qquad \text{by the induction hypothesis} \\ &\leq m\widehat{c}_s r \cdot ((\bar{\alpha}_{d,s}(m) - 1)^{s-2} + \bar{\alpha}_{d,s}(m)^{s-3}) \\ &\leq \widehat{c}_s rm\bar{\alpha}_{d,s}(m)^{s-2}. \end{aligned}$$

□

Let $\beta_s(m) := D_s(\alpha(m))$.

Corollary 3.17. *An $m \times n$ matrix M with at least $\beta_s(m)$ 1-entries in every column contains a doubled $(\lfloor (n-1)/(mc'_s) \rfloor, s)$-formation, where c'_s is a constant depending only on s.*

Proof. Let $c'_s = c_s 3^{s-3}$, where c_s is the constant from Corollary 3.16 and let $r = \lfloor (n-1)/(mc'_s) \rfloor$. If M did not contain a doubled (r,s)-formation, its number of columns would be, by Corollary 3.16 with $d = \alpha(m)$, at most

$$c_s r m \alpha_{\alpha(m)}(m)^{s-3} \leq r m c_s 3^{s-3} = \lfloor (n-1)/(mc'_s) \rfloor m c'_s < n. \qquad \square$$

A set S of Brs 1-entries forms a *B-fat (r,s)-formation* in M if there exists an s-partition of the rows and an r-tuple of columns each of which has B 1-entries of S in each interval. A matrix M is *B-fat (r,s)-formation-free* if it contains no B-fat (r,s)-formation.

We now prove a more precise version of Lemma 3.4.

Lemma 3.18. *For all positive integers m, n, s and B, an $m \times n$ matrix M with at least $2(\beta_s(m)+2)Bn$ 1-entries contains a B-fat $(\lfloor nB/(mc_s) \rfloor, s)$-formation, where c_s is a constant depending only on s.*

Proof. We transform the given matrix M to a matrix $\overline{M}$ with the same number of 1-entries in every column using the idea from the proof of Lemma 4.1 from [Niv10]. Let $v(q)$ be the number of 1-entries in a column q of M. In every column q, we split the 1-entries into chunks of consecutive $(\beta_s(m)+2)B$ 1's. The last less than $(\beta_s(m)+2)B$ 1's are discarded. Each of the chunks gets its own column with 1-entries in the rows where the 1-entries of the chunk lie. These columns form the matrix $\overline{M}$. Note that the order in which the columns are placed to $\overline{M}$ is not important. Because we discarded at most $(\beta_s(m)+2)Bn$ 1's and every column of $\overline{M}$ has exactly $(\beta_s(m)+2)B$ 1's, $\overline{M}$ has at least n columns. Observe that for every r and s if $\overline{M}$ contains a B-fat (r,s)-formation, then so does M.

We consider only the first n columns of $\overline{M}$. We also remove at most $B-1$ rows so as to have the number of rows divisible by B. We still have at least $(\beta_s(m)+1)B$ 1's in every column. In each column q, we select a set S of 1's such that none of them is among the first or the last $B-1$ 1's of the column q and there are at least $B-1$ 1's between every two 1's of S. We take S of size $\beta_s(m)$ and remove all the other 1's in q. The rows of $\overline{M}$ are now grouped into intervals of rows $\{iB+1, \ldots, (i+1)B\}$. By the choice of S, every column contains at most one 1-entry in every interval. We obtain $\overline{\overline{M}}$ by contracting each of the intervals of rows into a single row.

The matrix $\overline{\overline{M}}$ has $\lfloor m/B \rfloor$ rows and n columns, each of them having $\beta_s(m)$ 1's. It thus contains a doubled $(\lfloor (n-1)B/(mc'_s) \rfloor, s)$-formation by Corollary 3.17. By the choice of S, this implies that $\overline{M}$ and consequently M contain a B-fat $(\lfloor (n-1)B/(mc'_s) \rfloor, s)$-formation. $\square$

Proof of Lemma 3.4. Let $\zeta_s(m) = 2(\beta_s(m)+2)\max\{1, c_s\}$, where c_s is the constant from Lemma 3.18. Let M be an $m \times n$ matrix with at least $\zeta_s(m)Bn$ 1-entries. By Lemma 3.18, M contains a B-fat $(\lfloor nB/m \rfloor, s)$-formation. $\square$

3.2.3 Sets of permutations with bounded VC-dimension

In this section we prove Theorem 3.1. It will be more convenient for the proof to substitute the permutations by their corresponding permutation matrices. That is, we have a set $\mathcal{P}$ of n-permutation matrices and for every $(k+1)$-tuple $(a_1, \ldots, a_{k+1})$ of columns, there is a forbidden $(k+1)$-permutation matrix $S_{a_1,\ldots,a_{k+1}}$.

Let $M_{\mathcal{P}}$ be a $\{0,1\}$-matrix with 1-entries on the positions where at least one matrix from $\mathcal{P}$ has a 1-entry. Let $|M|$ denote the number of 1-entries in a $\{0,1\}$-matrix M and let $v(\mathcal{P}) = v(M_{\mathcal{P}}) = |M_{\mathcal{P}}|/n$. Similarly to Raz's proof of the exponential upper bound on $r_2(n)$ [Raz00], we will remove matrices from $\mathcal{P}$ until we decrease $v(\mathcal{P})$ below some threshold $T(n)$. When $v(\mathcal{P}) \le T(n)$, then $|\mathcal{P}| \le T(n)^n$ since the number of permutation matrices contained in $M_{\mathcal{P}}$ is bounded from above by the maximum of a product of n numbers with sum $v(\mathcal{P})n$.

Let $\gamma_k(n) = 2(k+1)!\zeta_{k+1}(n)$, where $\zeta_{k+1}(n)$ is the function from Lemma 3.4.

Lemma 3.19. *Let $\mathcal{P}$ be a set of n-permutation matrices with VC-dimension k such that $v(\mathcal{P}) \ge 2\gamma_k(n)$. Then there is a set $\mathcal{P}' \subset \mathcal{P}$ satisfying*

$$v(\mathcal{P}') \le v(\mathcal{P}) - \frac{v(\mathcal{P})^2}{\gamma_k^2(n)n}$$

$$|\mathcal{P}'| \ge \frac{|\mathcal{P}|}{2v(\mathcal{P})^k}.$$

Proof. Let $B := \lfloor v(\mathcal{P})/\zeta_{k+1}(n) \rfloor$. By Lemma 3.4, the matrix $M_{\mathcal{P}}$ contains a B-fat $(B, k+1)$-formation. Let C be the set of the B columns of the formation and let $\mathcal{R} = \{R_1, R_2, \ldots, R_{k+1}\}$ be the intervals of rows of the formation.

Consider some t-tuple $Q = \{q_1, \ldots, q_t\}$ of columns from C, where $t \le k+1$. Let $\mathcal{I}_Q$ be the set of all injective functions $I : Q \to \mathcal{R}$ assigning the intervals R_j to the columns q_i. We say that a permutation matrix P *obeys* $I \in \mathcal{I}_Q$ if for every $i \in \{1, 2, \ldots, t\}$, the 1-entry of P in the column q_i lies in some row of $I(q_i)$. For each $I \in \mathcal{I}_Q$ let $\mathcal{P}_I$ be the set of matrices $P \in \mathcal{P}$ that obey I. The t-tuple Q of columns is said to be *criss-crossed* if

$$\forall I \in \mathcal{I}_Q : |\mathcal{P}_I| \ge |\mathcal{P}|/v(\mathcal{P})^t.$$

Suppose that some $(k+1)$-tuple of columns from C is criss-crossed. Then every $(k+1)$-permutation appears as a restriction of some matrix from $\mathcal{P}$ on the criss-crossed $(k+1)$-tuple of columns. Hence the VC-dimension of $\mathcal{P}$ is at least $k+1$.

Consequently, there is some t such that $0 \le t \le k$ and the largest criss-crossed set Q of columns from C has size t. This means that for every column u outside Q, we can find an injective function $J_u \in \mathcal{I}_{Q\cup\{u\}}$ such that $|\mathcal{P}_{J_u}| < |\mathcal{P}|/v(\mathcal{P})^{t+1}$. On the other hand, if we restrict J_u on Q, the resulting function $I_u := J_u \restriction Q$ satisfies $|\mathcal{P}_{I_u}| \ge |\mathcal{P}|/v(\mathcal{P})^t$. To each choice of $u \in C \setminus Q$, we assign the function $I_u \in \mathcal{I}_Q$ and the interval $J_u(u)$ of rows. Some function $I \in \mathcal{I}_Q$ was then assigned to at least $(|C| - k)/|\mathcal{I}_Q|$ columns. Because $B \ge 4(k+1)!$, we have

$$\frac{|C| - k}{|\mathcal{I}_Q|} \ge \frac{B - k}{(k+1)!} \ge \frac{v(\mathcal{P})}{2\zeta_{k+1}(n)(k+1)!} \ge \frac{v(\mathcal{P})}{\gamma_k(n)}.$$

Let $\mathcal{T}_I$ be the set of some $\lceil v(\mathcal{P})/\gamma_k(n) \rceil$ columns that were assigned the function I. Because $v(\mathcal{P}) \ge 2\gamma_k(n)$, we have

$$\frac{v(\mathcal{P})}{\gamma_k(n)} \le |\mathcal{T}_I| \le \frac{2v(\mathcal{P})}{\gamma_k(n)} \le \frac{v(\mathcal{P})}{2}. \tag{3.1}$$

For each column $q_i \in Q$, we remove from M all 1-entries in the column q_i except those that lie in the rows of $I(q_i)$. This reduces the number of permutation matrices, but there are still at least $|\mathcal{P}|/v(\mathcal{P})^t$ of them. Then we remove from M the 1-entries in each column $u \in \mathcal{T}_I$ that lie in the set of rows $J_u(u)$. See Fig. 3.2. Thus we removed at least B 1-entries from each of these columns. The removed 1-entries of each of these columns decreased the number of permutation matrices by at most $|\mathcal{P}_{J_u}| \leq |\mathcal{P}|/v(\mathcal{P})^{t+1}$.

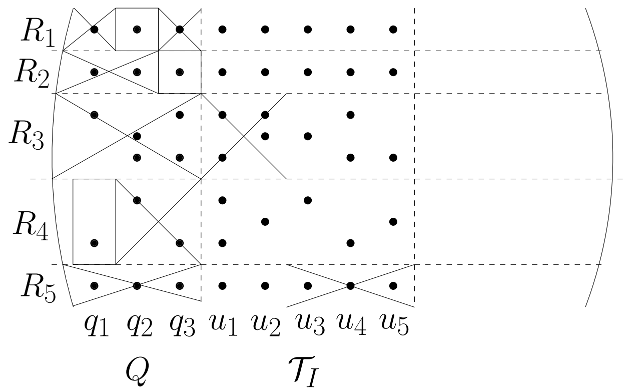


Figure 3.2: A criss-crossed set Q of 3 columns and a set $\mathcal{T}_I$ of columns $\{u_1, \ldots, u_5\}$, where $I(q_1) = R_4$, $I(q_2) = R_1$, $I(q_3) = R_2$, $J_{u_1}(u_1) = J_{u_2}(u_2) = R_3$ and $J_{u_3}(u_3) = J_{u_4}(u_4) = J_{u_5}(u_5) = R_5$. The 1-entries from the crossed rectangles are removed.

Let $\mathcal{P}' \subset \mathcal{P}$ be the set of permutation matrices containing none of the removed 1-entries. Using the bounds from Equation (3.1), we obtain

$$|M_{\mathcal{P}'}| \leq |M_{\mathcal{P}}| - B|\mathcal{T}_I| \leq nv(\mathcal{P}) - \frac{v(\mathcal{P})^2}{\gamma_k^2(n)},$$
$$|\mathcal{P}'| \geq \frac{|\mathcal{P}|}{v(\mathcal{P})^t} - \frac{|\mathcal{P}||\mathcal{T}_I|}{v(\mathcal{P})^{t+1}} \geq \frac{|\mathcal{P}|}{2v(\mathcal{P})^t} \geq \frac{|\mathcal{P}|}{2v(\mathcal{P})^k}. \qquad \square$$

Proof of Theorem 3.1. Let $\mathcal{P}$ be a set of permutation matrices with VC-dimension k. We will bound its size by iteratively applying Lemma 3.19. Let $\mathcal{P}_0 = \mathcal{P}$ and for $j \geq 1$ let $\mathcal{P}_j$ be the $\mathcal{P}'$ given by the lemma applied on $\mathcal{P}_{j-1}$.

The iterations are further grouped into phases. Let $\phi_0 := 0$. Phase i ends after the first iteration ϕ_i after which $v(\mathcal{P}_{\phi_i}) \leq v(\mathcal{P}_{\phi_{i-1}})/2$. Let $v_i := v(\mathcal{P}_{\phi_i})$. Then an iteration of phase i is applied on a set $\mathcal{P}$ of permutations satisfying

$$\frac{v_{i-1}}{2} \leq v(\mathcal{P}) \leq v_{i-1} \tag{3.2}$$

Further, let

$$T := \gamma_k^2(n) \log(\gamma_k(n)). \tag{3.3}$$

We end after the first phase l satisfying $v_l \leq 2T$. We thus have

$$|\mathcal{P}_{\phi_l}| \leq (2T)^n. \tag{3.4}$$

For every $i \geq 1$ we have

$$v_{l-i} \geq 2^i T. \tag{3.5}$$

We now count the number of iterations in phase i. By Lemma 3.19 and (3.2), each of these iterations decreases $v(\mathcal{P})$ by at least $v_{i-1}^2/(4\gamma_k^2(n)n)$. Therefore the phase ends after at most $\lceil 2\gamma_k^2(n)n/v_{i-1}\rceil \leq 3\gamma_k^2(n)n/v_{i-1}$ iterations. Consequently

$$\begin{aligned}
|\mathcal{P}_{\phi_{i-1}}| &\leq |\mathcal{P}_{\phi_i}| \cdot \left(2v_{i-1}^k\right)^{3\gamma_k^2(n)n/v_{i-1}} && \text{by Lemma 3.19 and (3.2)} \\
&\leq |\mathcal{P}_{\phi_i}| \cdot 2^{(1+k\log v_{i-1})3\gamma_k^2(n)n/v_{i-1}} \\
&\leq |\mathcal{P}_{\phi_i}| \cdot 2^{6k\gamma_k^2(n)n\log v_{i-1}/v_{i-1}}.
\end{aligned}$$

Then

$$|\mathcal{P}| = |\mathcal{P}_0| \leq |\mathcal{P}_{\phi_l}| \prod_{i=0}^{l-1} 2^{6k\gamma_k^2(n)n\log v_i/v_i}.$$

By (3.5) and since $\frac{\log(x)}{x}$ is decreasing on $[2T, \infty)$, we obtain

$$\begin{aligned}
|\mathcal{P}| &\leq |\mathcal{P}_{\phi_l}| 2^{6k\gamma_k^2(n)n\sum_{i=1}^{l}\log(2^iT)/(2^iT)} \\
&\leq (2T)^n \cdot 2^{6k\gamma_k^2(n)n(2+\log T)/T} && \text{by (3.4)} \\
&\leq \left(2\gamma_k^2(n)\log(\gamma_k(n))\right)^n \cdot 2^{30kn} && \text{by (3.3).}
\end{aligned}$$

Since $\gamma_k(n) \in O(\zeta_{k+1}(n))$, we have

$$\begin{aligned}
r_3(n) &\leq \left(O(\alpha(n)^4\log(\alpha(n)))\right)^n, \\
r_4(n) &\leq 2^{n\cdot(2\alpha(n)+3\log(\alpha(n))+O(1))}, \\
r_{2t+2}(n) &\leq 2^{n\cdot\left((2/t!)\alpha(n)^t+O(\alpha(n)^{t-1})\right)} && \text{for } t \geq 1 \text{ and} \\
r_{2t+3}(n) &\leq 2^{n\cdot\left((2/t!)\alpha(n)^t\log(\alpha(n))+O(\alpha(n)^t)\right)} && \text{for } t \geq 1.
\end{aligned}$$

□

Remark. Let an n-word be a word of length n over the alphabet $[n]$. We say that a set $\mathcal{W}$ of n-words has *VC-dimension with respect to permutations* (abbreviated as *pVC-dimension*) k if k is the largest integer such that the set of restrictions of the words in $\mathcal{F}$ to some k-tuple of positions from $[n]$ contains all k-permutations. Let $r'_k(n)$ be the size of the largest set of n-words with pVC-dimension k. The proofs of this section never use the fact that the matrices in $\mathcal{P}$ have exactly one 1-entry in every row. Thus the upper bound from Theorem 3.1 also holds with $r'_k(n)$ in place of $r_k(n)$.

3.3 Lower bounds

3.3.1 Matrices from sequences

Let DS_s be the $s \times 2$ matrix with 1-entry in the ith row and jth column exactly when $i + j$ is odd. For example

$$DS_4 = \begin{pmatrix} & \bullet \\ \bullet & \\ & \bullet \\ \bullet & \end{pmatrix} \qquad \text{and} \qquad DS_5 = \begin{pmatrix} & \bullet \\ \bullet & \\ & \bullet \\ \bullet & \\ & \bullet \end{pmatrix}.$$

Based on a construction of Davenport–Schinzel sequences of order 3 and length $\Omega(n\alpha(n))$ by Hart and Sharir [HS86], Füredi and Hajnal [FH92] constructed $n\times n$

DS_4-avoiding matrices A_n with $\Omega(n\alpha(n))$ 1-entries. We will use a different construction of DS(s)-sequences of orders $s = 3$ and all even $s \geq 4$ by Nivasch [Niv10] that together with the following transcription will provide us with DS_{s+1}-avoiding matrices with the additional property of having the same number of 1-entries in every column.

Let S be a sequence over n symbols that can be partitioned into m blocks. Recall that each block contains only distinct symbols. We number the symbols $1, \ldots, n$ in the increasing order of their first appearance. The *sequence→matrix transcription of S,* SMT(S), is the $m \times n$ $\{0, 1\}$-matrix with a 1-entry in the ith row and jth column exactly if the ith block in the sequence contains the symbol j.

Observation 3.20. ([FH92]) *If S is a sequence avoiding the alternating pattern $aba\ldots$ of length $s+2$, then* SMT(S) *avoids DS_{s+1}.*

Proof. If SMT(S) contains DS_{s+1}, then S contains the alternating sequence $ba\ldots$ of length $s + 1$ for some $a < b$. By the numbering of the symbols, the symbol a appears before the first occurrence of b, therefore S contains $aba\ldots$ of length $s + 2$ and thus S is not a DS(s)-sequence. □

Lemma 3.21. *For every n there exists an $n \times n$ DS_4-avoiding matrix M_n with at least $2\alpha(n) - O(1)$ 1-entries in every column.*

Proof. Let $A_d(x)$ be the dth function of the Ackermann hierarchy. We refer the reader to Nivasch's paper [Niv10] for the definition. Let $A(x) = A_x(3)$ be the Ackermann function.

In Section 6 of [Niv10], Nivasch constructs for every $d, m \geq 1$ an *ababa*-free sequence $Z_d(m)$. We use the sequences $Z'_d = Z_d(8d+4)$ which have the following properties:

- Each symbol appears exactly $2d + 1$ times.
- The sequence can be decomposed into blocks of average length at least $4d+2$ (by [Niv10, Lemma 6.2]).
- The number N_d of symbols of the sequence is at most $A_d(8d + 4 + c)$ (by [Niv10, Lemma 6.2]), where c is an absolute constant.

Let M_d be the number of blocks of Z'_d. By counting the length of Z'_d in two ways, $(2d + 1)N_d \geq (4d + 2)M_d$ and thus $N_d \geq 2M_d$. By the analysis before Equation (35) in [Niv10], there is some d_0 such that for $d \geq d_0$ we have

$$N_d \leq A_d(8d + 4 + c) \leq A_d(A(d + 1)) = A(d + 2) \quad \text{and so}$$
$$\alpha(N_d) \leq d + 2.$$

Then SMT(Z'_d) is an $M_d \times N_d$ matrix with $2d + 1 \geq 2\alpha(N_d) - 3$ 1-entries in every column. By Observation 3.20, SMT(Z'_d) avoids DS_4. We construct the matrix M_{N_d} by adding empty rows to SMT(Z'_d).

For values $n \geq N_{d_0}$ different from N_d, we proceed similarly to the method in Section 6 of [Niv10]: We consider the largest N_d smaller than n and take $\lceil n/N_d \rceil$ copies of SMT(Z'_d). We place the copies into a single matrix so that each copy has its own set of consecutive rows and columns. After removing at most half of the columns, we obtain a matrix with exactly n columns and at most n rows. The matrix has at least $2d + 1 \geq 2\alpha(N_{d+1}) - 5 \geq 2\alpha(n) - 5$ 1-entries in every column. The construction of M_n is then finished by adding empty rows to obtain a square matrix. $\square$

Lemma 3.22. *For every $t \geq 1$ and n there exists an $n \times n$ DS_{2t+3}-avoiding matrix with at least $2^{(1/t!)\alpha(n)^t - O(\alpha(n)^{t-1})}$ 1-entries in every column. In particular,*

$$\mathrm{ex}_{DS_{2t+3}}(n) \geq n2^{(1/t!)\alpha(n)^t - O(\alpha(n)^{t-1})}.$$

Proof. Let $s := 2t + 2$. Since s is even and $s \geq 4$, we can use Nivasch's construction [Niv10, Section 7] of DS(s)-sequences $S_k^s(m)$ with parameters $k, m \geq 2$. Let $\mu_s(k) := 2^{\binom{k}{(s-2)/2}}$. We use the sequences $S'_{s,k} = S_k^s(2\mu_s(k))$, which have the following properties:

- Each symbol of $S'_{s,k}$ appears exactly $\mu_s(k)$ times (by [Niv10, Equation (47)]).
- The sequence can be decomposed into blocks of length $2\mu_s(k)$.
- For every $k \geq k_0(s)$, where $k_0(s)$ is a properly chosen constant, the number $N_{s,k}$ of symbols of the sequence satisfies $\alpha(N_{s,k}) \leq k + 3$ (by [Niv10, Equations (50), (51)] and analysis similar to the one in the proof of Lemma 3.21).

Let $M_{s,k}$ be the number of blocks of $S'_{s,k}$. It satisfies $2M_{s,k} \leq N_{s,k}$. The matrix SMT($S'_{s,k}$) is a DS_{s+1}-avoiding $M_{s,k} \times N_{s,k}$ matrix with at least $\mu_{s,k}$ 1-entries in every column. For every $n \geq N_{s,k_0(s)}$ we take the largest k such that $n \geq N_{s,k}$ and proceed in the same way as in the proof of Lemma 3.21 with SMT($S'_{s,k}$) in the place of SMT(Z'_d). We have

$$k \geq \alpha(N_{s,k+1}) - 4 \geq \alpha(n) - 4$$

and so the number of 1-entries in every column of the resulting matrix is

$$\mu_s(k) = 2^{\binom{k}{(s-2)/2}} = 2^{\binom{k}{t}} \geq 2^{(1/t!)k^t - O(k^{t-1})} \geq 2^{(1/t!)\alpha(n)^t - O(\alpha(n)^{t-1})}. \qquad \square$$

Remark. We could also use the construction of DS_4-avoiding matrices with $\Omega(n\alpha(n))$ 1-entries by Füredi and Hajnal [FH92]. The matrices do not have the same number of 1-entries in every column, but it can be shown that every column has at most a constant multiple of the average number of 1-entries per column. This would be enough for our purposes. The base case of the inductive construction in [FH92] needs a small fix. The matrices $M(s, 1)$ and $M(1, s)$ do not satisfy conditions imposed on them. This can be fixed for example by taking $\begin{pmatrix}\bullet & \\ & \bullet\end{pmatrix}$ for $M(s, 1)$ and the matrix with the leftmost column full of 1-entries and with no 1-entries in the other columns for $M(1, s)$.

3.3.2 Numbers of 1-entries in matrices

A matrix is *k-full* if some k-tuple of its columns contains every k-permutation matrix. The *fullness* of a matrix A is the largest k such that A is k-full. In this section we show a lower bound on the maximum number $p_k(n)$ of 1-entries in an $n \times n$ matrix with fullness k. This is achieved by showing that a k-full matrix contains the matrix DS_k and applying the results from Section 3.3.1. We prove a slightly stronger statement that will be used in the next section.

Let $J_2 := \left(\begin{smallmatrix} & \bullet \\ \bullet & \end{smallmatrix}\right)$. For an l-permutation matrix P, we define the *J_2-expansion of P*, P^{J_2}, to be the $2l \times 2l$ permutation matrix created by substituting every 1-entry of P by J_2 and every 0-entry by a 2×2 block full of 0-entries.

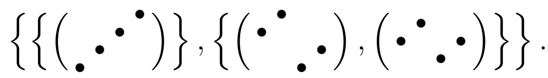
A pair of rows $2i$, $2i+1$ of P^{J_2} will be called *contractible* if the 1-entry in row $2i$ is to the left of the 1-entry in row $2i+1$. That is, when $\pi^{-1}(i) < \pi^{-1}(i+1)$, where π is the permutation corresponding to P. To *contract* a pair of rows means to replace them by a single row with 1-entries in the columns where at least one of the two original rows had a 1-entry.

Let an (n, m)-word be a word from $[m]^{[n]}$. A *word matrix* is a $\{0, 1\}$-matrix with exactly one 1-entry in every column. Assigning to a word f a word matrix G_f with $G_f(i, j) = 1 \Leftrightarrow f(j) = i$ provides a bijection between (n, m)-words and $m \times n$ word matrices.

The set of *J_2-expansion flattenings* of P is the set $\mathcal{F}(P^{J_2})$ of word matrices that can be obtained from P^{J_2} by contracting some pairs of contractible rows. Let $\mathcal{P}_l$ be the set of l-permutation matrices and let

$$\Phi(l) := \{\mathcal{F}(P^{J_2}) : P \in \mathcal{P}_l\}.$$

For example

$\Phi(2) =$ 

Lemma 3.23. *If an $n \times 2l$ matrix A contains one matrix from $\mathcal{F}(P^{J_2})$ for every l-permutation matrix P, then A contains an occurrence of DS_{2l} on columns $\{2i - 1, 2i\}$ for some $i \in [l]$.*

Proof. We proceed by induction on l. The case $l = 1$ is trivial since $\Phi(l) = \{\{J_2\}\} = \{\{DS_2\}\}$.

The *ith pair of columns* of A is the pair of columns $\{2i - 1, 2i\}$. For each $i \le l$ let h_i be the smallest number such that the ith pair of columns of A contains J_2 on a subset of rows $\{1 \dots h_i\}$. Let t be the largest number satisfying $\forall i\ h_i \le h_t$. Let $A^{\setminus t}$ be the $(n - h_t + 1) \times 2(l - 1)$ matrix obtained from A by removing the columns of the tth pair, removing the top $h_t - 1$ rows and then changing all 1-entries among the first $2(t - 1)$ entries in the first row to 0's. See Fig. 3.3

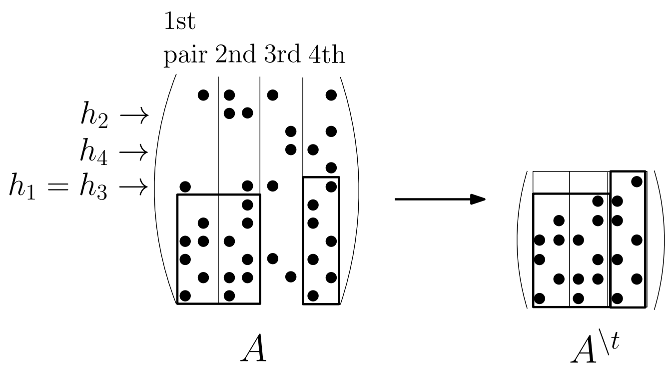
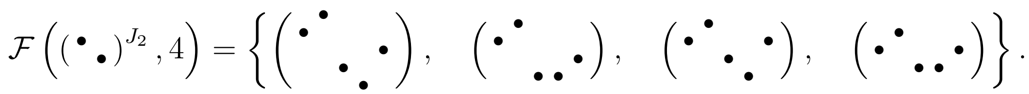


Figure 3.3: Induction step in the proof of Lemma 3.23. In this example $t = 3$.

For every P with the topmost 1-entry in column t, A contains an occurrence of some $F \in \mathcal{F}(P^{J_2})$, that uses the two 1-entries of the topmost occurrence of J_2 on the tth pair of columns. These occurrences induce an occurrence of some matrix from every set $\mathcal{F} \in \Phi(l-1)$ in $A^{\setminus t}$. By the induction hypothesis, $A^{\setminus t}$ contains $DS_{2(l-1)}$ on some ith pair of columns. By the choice of t, this occurrence of $DS_{2(l-1)}$ in A does not use any of the rows $\{1 \dots h_i\}$. Thus we obtain an occurrence of DS_{2l} in A. □

For an l-permutation matrix P and $i \leq 2l+1$ we define $P^{J_2}(i)$ to be the $(2l+1)$-permutation matrix that becomes P^{J_2} after removing the lowest row and column i. Then $\mathcal{F}(P^{J_2}, i)$ is the set of word matrices that can be obtained from $P^{J_2}(i)$ by contracting some pairs of contractible rows. For example



Let

$$\Phi(l, i) := \{\mathcal{F}(P^{J_2}, i) : P \in \mathcal{P}_l\}.$$

Lemma 3.24. *Let A be an $n \times (2l+1)$ matrix and let A' be the matrix obtained from A by removing the last 1-entry from each column. If A' contains one matrix from $\mathcal{F}(P^{J_2}, i)$ for every l-permutation matrix P and every $i \in [2l+1]$, then A contains DS_{2l+1}.*

Proof. For each $i \leq 2l+1$ let d_i be the row number of the lowest 1-entry in the ith column of A'. Let t be any of the rows satisfying $\forall i \; d_i \geq d_t$. Let $A^{\setminus t}$ be the $d_t \times 2l$ matrix obtained from A by removing the tth column and all rows below the d_tth row. Then $A^{\setminus t}$ contains one matrix from every $\mathcal{F} \in \Phi(l)$, therefore by Lemma 3.23 the matrix $A^{\setminus t}$ contains an occurrence of DS_{2l}. By the choice of t, the matrix A contains DS_{2l+1}. □

Corollary 3.25. *For every $k \geq 1$*

$$p_k(n) \geq \mathrm{ex}_{DS_{k+1}}(n).$$

Proof. When k is even, the result follows from Lemma 3.24, since for every l-permutation matrix P and for every $i \in [2l+1]$, the set $\mathcal{F}(P^{J_2}, i)$ contains some $(2l+1)$-permutation matrix, namely the matrix without any row contractions. The result for k odd follows from Lemma 3.23. □

The row contractions did not play any role in the proof of Corollary 3.25, but they will play a role in Section 3.3.3 below.

Corollary 3.26. *We have*

$$p_3(n) \geq 2n\alpha(n) - O(n),$$
$$p_k(n) \geq n2^{(1/t!)\alpha(n)^t - O(\alpha(n)^{t-1})} \quad \textit{for } k \geq 4,$$

where $t := \lfloor (k-2)/2 \rfloor$.

Proof. The lower bound for $k = 3$ is by Lemma 3.21 and Corollary 3.25 and from Lemma 3.22 and Corollary 3.25 when k is even and $k > 3$. When k is odd and $k \geq 5$, we use $p_k(n) \geq p_{k-1}(n)$. □

3.3.3 Sets of permutations

Proof of Theorem 3.2. Given k and n, we take the DS_{k+1}-avoiding $n \times n$ matrix $A_{k,n}$ from Lemma 3.21 if $k = 3$ or from Lemma 3.22 if $k \geq 4$ is even. Let $\rho_k(n)$ be the number of 1-entries that $A_{k,n}$ has in every column, that is $\rho_2(n) = 2\alpha(n) - O(1)$ and for $t \geq 1$ $\rho_{2t+2}(n) = 2^{(1/t!)\alpha(n)^t - O(\alpha(n)^{t-1})}$.

From $A_{k,n}$ we construct a set of $\rho_k(n)^n$ $n \times n$ word matrices by choosing some 1-entry from each column. Then we remove all empty rows, which can make some originally different word matrices identical. However, the resulting set $\mathcal{H}$ has size at least $\rho_k(n)^n/2^n$ as there are at most 2^n distinct ways to enlarge a word matrix by adding empty rows to a matrix with n rows.

The last step is inflating the rows of the word matrices in $\mathcal{H}$ into identity matrices to obtain a set $\mathcal{Q}$ of n-permutation matrices. That is, for every $H \in \mathcal{H}$, we order the 1-entries primarily by the rows from top to bottom and secondarily from left to right. The permutation matrix Q has 1-entries on those positions (i, j) such that H has its ith 1-entry in column j. The reverse process consists of contracting intervals of rows of a permutation matrix Q and we have at most 2^n possibilities how to choose the intervals. Thus every permutation matrix can be created by expanding at most 2^n different word matrices. The size of the set $\mathcal{Q}$ is

$$|\mathcal{Q}| \geq \frac{\rho_k(n)^n}{2^n 2^n} = \left(\frac{\rho_k(n)}{4}\right)^n.$$

It remains to show that the VC-dimension of $\mathcal{Q}$ is at most $(k+1)$. We assume for contradiction that for some $(k+1)$-tuple C of columns and every $(k+1)$-permutation matrix R there exists $Q \in \mathcal{Q}$ that contains R on C.

Consider some permutation matrix $Q \in \mathcal{Q}$ and let $H \in \mathcal{H}$ be the word matrix from which Q was created. The matrix H can thus be constructed from Q by contracting some intervals of rows such that the restriction of Q on each of these intervals of rows is the identity matrix. So the only change that these contractions can make on an occurrence of P^{J_2} in Q is that some pairs of its contractible rows can be contracted. Thus an occurrence of P^{J_2} in Q on the set C of columns can only be created from an occurrence of some $F \in \mathcal{F}(P^{J_2})$ on C in H and in $A_{k,n}$ as well. Similarly, an occurrence of $P^{J_2}(i)$ on C in Q implies an occurrence of some matrix from $\mathcal{F}(P^{J_2}, i)$ on C in H and $A_{k,n}$. See Fig. 3.4.

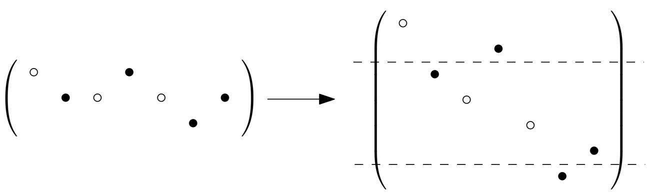


Figure 3.4: Expansion of an occurrence of a matrix from $\mathcal{F}(P^{J_2})$.

Therefore for $k \geq 4$ even, for every $(k/2)$-permutation matrix P and every $i \in [k+1]$, some matrix from $\mathcal{F}(P^{J_2}, i)$ occurs on C in $A_{k,n}$. Thus, by Lemma 3.24, $A_{k,n}$ contains DS_{k+1}, a contradiction. Similarly if $k = 3$, we get a contradiction by Lemma 3.23. □

4. Reverse-free sets of permutations

4.1 Introduction

A word is *repetition-free* if it contains at most one occurrence of each symbol. The set of all repetition-free words of length k over $[n]$ is $[n]_{(k)}$. Notice that when $n = k$, the set $[n]_{(n)}$ is the set S_n of permutations on n elements . A *code* $\mathcal{F}$ of length k is a subset of $[n]^k$. The *size* of $\mathcal{F}$ is the number of words in $\mathcal{F}$. Codes are usually defined to be sets of words that in some sense significantly differ from each other in order to be distinguishable when transmitted over a noisy channel. We study reverse-free codes introduced by Füredi, Kantor, Monti and Sinaimeri [FKMS10]. Two words w and x have a *reverse* if for some pair (i, j) of positions, we have $w_i \neq w_j$, $w_i = x_j$ and $w_j = x_i$. If w and x do not have a reverse, they are *reverse-free*. A code is *reverse-free* if its words are pairwise reverse-free. Let $\overline{F}(n, k)$ be the size of the largest reverse-free code over $[n]$ of length k. Let $F(n, k)$ be the size of the largest reverse-free code over $[n]$ of length k containing only repetition-free words. Let

$$f(k) \stackrel{def}{=} \lim_{n\to\infty} \frac{F(n, k)}{k!\binom{n}{k}}.$$

The limit exists for every $k \geq 1$ [FKMS10]. We will use the following equivalent definitions of the limit:

$$f(k) = \lim_{n\to\infty} \frac{F(n, k)}{n^k} = \lim_{n\to\infty} \frac{\overline{F}(n, k)}{n^k}.$$

The first equality follows from the fact that $\lim_{n\to\infty} \binom{n}{k} k! n^{-k} = 1$. The second equality is a consequence of the observation that for every fixed k, we have $F(n, k) \leq \overline{F}(n, k) \leq F(n, k) + O(n^{k-1})$ [FKMS10]. The only exact values of the limit known are $f(1) = 1$, $f(2) = 1/2$ and $f(3) = 5/24$ [FKMS10].

We tighten the bounds on the maximum size of reverse-free codes of length greater or equal to the size of the alphabet.

Theorem 4.1. *For every $n \geq k$, we have*

$$n^k k^{-k/2-O(k/\log k)} \leq F(n, k) \leq \overline{F}(n, k) \leq n^k k^{-k/2+O(k/\log k)}.$$

The first inequality is proven in Section 4.2 as Corollary 4.7 and the last inequality is proven as Claim 4.11 in Section 4.3. As an immediate consequence, we obtain the following bounds for permutation codes:

$$n^{n/2-O(n/\log n)} \leq F(n, n) \leq \overline{F}(n, n) \leq n^{n/2+O(n/\log n)}$$

and for the limit for codes of fixed length k:

$$f(k) \in k^{-k/2+O(k/\log k)}.$$

A set of words is *full of flips* if each two words from the set have a reverse. Let $\overline{G}(n, k)$ be the size of the largest code full of flips with elements in $[n]^k$. Let $G(n, k)$ be the size of the largest code full of flips with elements in $[n]_{(k)}$. By $G(n, n)F(n, n) \leq n!$ [FKMS10], we obtain the following corollary.

Corollary 4.2. *The size of a set of permutations full of flips is at most*

$$G(n, n) \leq n^{n/2+O(n/\log n)}.$$

A position of an entry of a matrix is represented by a pair (r, c) of the row number r and the column number c. All logarithms in this chapter are of base 2.

4.2 Lower Bound

We consider matrices avoiding the 2×2 matrix $S = \begin{pmatrix} \bullet & \bullet \\ \bullet & \bullet \end{pmatrix}$. Maximizing the number of 1's in a matrix avoiding S is closely related to maximizing the number of edges in an n-vertex graph without a 4-cycle as a subgraph [FH92]. The maximum number of edges in a 4-cycle-free graph is known precisely for infinitely many values of n [Für96]. We will use a classical construction of a bipartite 4-cycle-free graph (see for example the book of Matoušek and Nešetřil [MN98]). We reproduce the construction here in the matrix setting since we need some of its additional properties.

The construction of a matrix avoiding S builds the matrix using a finite projective plane. Let X be a finite set and let $\mathcal{L}$ be a family of subsets of X. The set system $(X, \mathcal{L})$ is a *finite projective plane* if

(P0) There is a 4-tuple F of elements of X such that $|F \cap L| \leq 2$ for every $L \in \mathcal{L}$.

(P1) For every $L_1, L_2 \in \mathcal{L}$, $|L_1 \cap L_2| = 1$.

(P2) For every $x, y \in X$ there exists exactly one $L \in \mathcal{L}$ containing both x and y.

For every finite projective plane, we can find a number r, called the *order of the projective plane*, satisfying:

(P3) For every $L \in \mathcal{L}$, $|L| = r + 1$.

(P4) Every $x \in X$ is contained in exactly $r + 1$ sets from $\mathcal{L}$.

(P5) We have $|X| = |\mathcal{L}| = r^2 + r + 1$. This value is the *size* of the projective plane.

It is known that for every number r that is a power of a prime number, we can find a finite projective plane of order r [MN98].

Claim 4.3. *If n is of the form $r^2 + r + 1$, where r is a power of a prime, then*

$$F(n, n) \geq n^{n/2-O(n/\log n)}.$$

Proof. We fix a projective plane $(X, \mathcal{L})$ of size n. We order the elements of X and the sets of $\mathcal{L}$ arbitrarily. The *incidence matrix* of a finite projective plane of size n is the $n \times n$ matrix A with 1 on position (i, j) exactly if the i-th set of $\mathcal{L}$ contains the j-th element of X. Let A be the incidence matrix of $(X, \mathcal{L})$.

Let $\mathcal{P}$ be the set of n-permutation matrices contained in A and let Π be the set of n-permutations matched to the matrices from $\mathcal{P}$. By (P3) and (P4), A has exactly $r+1$ 1's in every row and every column. Thus by the van der Waerden conjecture proved independently by Falikman [Fal81] and Egorychev [Ego81],

$$|\Pi| = |\mathcal{P}| \geq \left(\frac{r+1}{n}\right)^n n! \geq \left(\frac{r+1}{e}\right)^n \geq \left(\frac{n^{1/2}}{e}\right)^n \geq n^{n/2 - O(n/\log n)}.$$

We claim that the set Π is pairwise reverse-free. For contradiction, we take $\pi \in \Pi$ and $\rho \in \Pi$ with a reverse on positions i and j. That is, $\pi_i = k$, $\pi_j = l$, $\rho_i = l$, $\rho_j = k$ for some k and l. Since P_π and P_ρ are contained in A, this implies that A contains the matrix S on rows k, l and columns i and j; a contradiction with (P1). $\square$

By the prime number theorem, the gaps between two consecutive prime numbers in proportion to the primes tend to zero. There has been a significant progress in tightening the gap between two consecutive primes. Most recent is the following result of Baker, Harman and Pintz [BHP01].

Theorem 4.4 (Baker, Harman, Pintz, 2001). *For every large enough n, the interval $[n - n^{0.525}, n]$ contains a prime number.*

Lemma 4.5. *For every n,*

$$F(n,n) \geq n^{n/2 - O(n/\log n)}.$$

Proof. For an arbitrary n we take the largest n' smaller than n and expressible as $p^2 + p + 1$ for some prime p. The interval $[n^{1/2} - 1 - n^{0.525/2}, n^{1/2} - 1]$ contains a suitable prime number p. Thus

$$p \geq n^{1/2} - 1 - n^{0.525/2} \quad \text{and}$$
$$n' \geq n - O(n^{1.525/2}).$$

We take the set Π' of $(n')^{n'/2 - O(n'/\log n')}$ pairwise reverse-free n'-permutations from Claim 4.3. We append the sequence $(n'+1, n'+2, \ldots, n)$ to the end of each $\pi' \in \Pi'$. Let the resulting set of n-permutations be Π. The set Π of permutations is pairwise reverse-free and has size at least $n^{n/2 - O(n/\log n)}$. $\square$

Lemma 4.6. *For every $n \geq k$,*

$$F(n,k) \geq \left\lfloor \frac{n}{k} \right\rfloor^k F(k,k).$$

Proof. Let Π be a reverse-free set of k-permutations of size $F(k,k)$. Given a word $u = (u_1, u_2, \ldots, u_k) \in [n]_{(k)}$, we call the word $(u_1 \bmod k, u_2 \bmod k, \ldots, u_k \bmod k)$ the *compression* of u. Let $\mathcal{F}$ be a set of all the words in $[n]_{(k)}$ whose compression is in Π. The size of $\mathcal{F}$ is at least $\left\lfloor \frac{n}{k} \right\rfloor^k |\Pi|$. It remains to show that $\mathcal{F}$ is reverse-free. For contradiction, assume that some pair of words $(u_1, \ldots, u_k)$ and $(v_1, \ldots, v_k)$ has a reverse on the pair (i,j) of positions. That is, $u_i = v_j$ and $u_j = v_i$ and, in particular, $u_i \bmod k = v_j \bmod k$ and $u_j \bmod k = v_i \bmod k$. Because the compression of u is a permutation, $u_i \bmod k \neq u_j \bmod k$. This is a contradiction, because the compressions of u and v are in the reverse-free set Π. $\square$

Corollary 4.7. *For every $n \geq k$,*

$$F(n,k) \geq n^k k^{-k/2-O(k/\log k)}.$$

Proof. Since $n \geq k$, we have $\lfloor n/k \rfloor \geq n/(2k)$. Therefore, by Lemmas 4.5 and 4.6

$$F(n,k) \geq \left(\frac{n}{2k}\right)^k F(k,k) \geq \frac{n^k}{(2k)^k} k^{k/2-O(k/\log k)} \geq n^k k^{-k/2-O(k/\log k)}.$$

□

4.3 Upper Bound

We use a result claiming that a matrix with many 1-entries contains many occurrences of the matrix $S = \begin{pmatrix}\bullet & \bullet \\ \bullet & \bullet\end{pmatrix}$. This corresponds to counting the occurrences of $K_{2,2}$ in a bipartite graph. Erdős and Simonovits [ES69] proved that an arbitrary graph G with $e(G)$ edges and $v(G)$ vertices contains at least $e(G)^4/(2v(G)^4) - e(G)^2/(2v(G))$ copies of $K_{2,2}$. Sidorenko [Sid91] proves a general result that also gives a lower bound on the number of occurrences of $K_{2,2}$ (and several other bipartite graphs) in a graph with many edges. It follows from [Sid91, Condition B] that a bipartite graph with parts of size n and k contains $e(G)^4/(4n^2k^2) - O(nk^2 + n^2k)$ copies of $K_{2,2}$.

Neither of these results is applicable in cases when k is much smaller than n and the number of edges is of the order $O(nk^{1/2})$. Thus, we follow the approach used in the above mentioned papers to prove the following lemma which gives a more precise bound in such cases. This approach appeared already in 1964 in a proof of a similar result of Erdős and Moon [EM64].

Lemma 4.8. *Let k and n be integers such that $n \geq k \geq 1$ and let m be a real number from the closed interval $[1, k^{1/2}]$. Let A be an $n \times k$ $\{0,1\}$-matrix with at least $mnk^{1/2}$ 1-entries. The number of occurrences of S in A is at least*

$$\frac{n^2(m^2-1)^2}{4} - m^3 nk^{1/2}.$$

Proof. If $k = 1$, then A has at most n 1's and so $m = 1$ and the claim is trivially satisfied. So we assume that $k \geq 2$. We also assume that A has no empty columns. If A has empty columns, we remove them and use the claim for the matrix with no empty columns. Since the removal decreases k, keeps n the same and increases m, we obtain the desired estimate.

We first count the number q of pairs of 1's that are in the same column. Let d_i be the number of 1's in the i-th column of A. We have

$$q = \sum_{\{i: d_i \geq 2\}} \binom{d_i}{2} \geq \sum_{i=1}^{k} \frac{(d_i-1)^2}{2}.$$

Let d be the average number of 1's in a column, that is,

$$d \stackrel{\text{def}}{=} \frac{\sum_{i=1}^{k} d_i}{k} \geq \frac{mnk^{1/2}}{k} = mnk^{-1/2}.$$

By the convexity of the function $f(x) = (x-1)^2/2$, we have

$$q \geq k\frac{(d-1)^2}{2} \geq \frac{k}{2}\left(mnk^{-1/2}-1\right)^2 \geq \frac{(mn-k^{1/2})^2}{2} \geq \frac{m^2n^2}{2} - mnk^{1/2}.$$

Let $r_{i,j}$ be the number of columns that have a 1-entry in rows i and j. Let $\mathcal{R}$ be the set of pairs $\{i,j\}$ of column indices satisfying $1 \leq i < j \leq n$ and $r_{i,j} \geq 2$.

First, we consider the case $q \leq n^2/2$. From the estimate $q \geq (mn - k^{1/2})^2/2$, we obtain $m \leq 1 + k^{1/2}/n$. Therefore $n^2(m^2-1)^2 \in O(k)$ and the result holds trivially, because $k \leq n$.

Now, we assume $q > n^2/2$, which implies $|\mathcal{R}| > 0$. By double counting,

$$q = \sum_{1 \leq i<j \leq n} r_{i,j} \leq \sum_{\{i,j\}\in\mathcal{R}} r_{i,j} + \binom{n}{2} - |\mathcal{R}|.$$

Let

$$r \stackrel{\text{def}}{=} \frac{\sum_{\{i,j\}\in\mathcal{R}} r_{i,j}}{|\mathcal{R}|} \geq \frac{q - \left(\binom{n}{2} - |\mathcal{R}|\right)}{|\mathcal{R}|} = \frac{q - \binom{n}{2}}{|\mathcal{R}|} + 1.$$

Let s be the number of occurrences of S in A, that is, $s = \sum_{\{i,j\}\in\mathcal{R}} \binom{r_{i,j}}{2}$. By the convexity of $f(x) = (x-1)^2/2$ and since $r > 1$, we have

$$\begin{aligned} s &\geq |\mathcal{R}|\frac{(r-1)^2}{2} \geq \frac{|\mathcal{R}|}{2}\left(\frac{q-\binom{n}{2}}{|\mathcal{R}|}\right)^2 \\ &\geq \frac{\left(m^2n^2/2 - mnk^{1/2} - n^2/2\right)^2}{2|\mathcal{R}|} \\ &\geq \frac{\left(n^2(m^2-1)/2 - mnk^{1/2}\right)^2}{n^2} \\ &\geq \frac{n^2(m^2-1)^2}{4} - m^3nk^{1/2}. \end{aligned}$$

□

We first give some definitions and outline the proof of the upper bound in Theorem 4.1 without mentioning precise values used. We use a modification of a method of Raz [Raz00], that was used in Chapter 3 for proving upper bounds in another extremal problem on sets of permutations [Raz00, CK12].

An $n \times k$ *word matrix* is a $n \times k$ matrix with exactly one 1-entry in every column. An (n,k)-*word* is a word of length k over the alphabet $[n]$. The following is a bijection between the set of (n,k)-words and the set of $n \times k$ word matrices. A word u is matched with the matrix U having 1 on position (i,j) exactly if $u_j = i$. A set $\mathcal{U}$ of $n \times k$ word matrices is *reverse-free* if the set of corresponding words is reverse-free.

Given a set $\mathcal{U}$ of $n \times k$ word matrices, we let the *overall matrix* $A_{\mathcal{U}}$ be the $n \times k$ matrix having 1-entries on those positions where at least one matrix of $\mathcal{U}$ has a 1-entry. The basic idea is to design a procedure that shrinks the set $\mathcal{U}$ in order to decrease the number of 1's in the overall matrix. When the overall matrix has few 1's, we use a trivial estimate on the size of what remained in $\mathcal{U}$. By analyzing the procedure, we then deduce that the original size of $\mathcal{U}$ was small.

The shrinking procedure uses the result of Lemma 4.8 applied on the overall matrix. Assume that the overall matrix contains S on the intersection of rows r_1 and r_2 and columns c_1 and c_2. Let an *avoided pair* be a pair of 1-entries of the overall matrix that do not appear together in any matrix in $\mathcal{U}$. By the reverse-free property of $\mathcal{U}$, we know that at least one of the two pairs $\{(r_1, c_1), (r_2, c_2)\}$ and $\{(r_1, c_2), (r_2, c_1)\}$ is avoided. When the overall matrix contains many occurrences of S, we find a 1-entry (r, c) occurring in many avoided pairs. If the 1-entry (r, c) is not present in enough matrices from $\mathcal{U}$, we remove from $\mathcal{U}$ all the matrices containing (r, c), thus removing (r, c) from the overall matrix. Otherwise, we keep only the matrices that contain (r, c), thus removing all the matrices containing any of the 1-entries that appear in some avoided pair together with (r, c).

Given a reverse-free set $\mathcal{U}$ of $n \times k$ word matrices, let $A_{\mathcal{U}}$ be the overall matrix of $\mathcal{U}$. Let the *weight* $|A_{\mathcal{U}}|$ of the overall matrix be the number of its 1-entries. The *density* of the overall matrix is $m_{\mathcal{U}} = |A_{\mathcal{U}}|/(nk^{1/2})$. The 1-entry of the overall matrix $A_{\mathcal{U}}$ on the position (r, c) is *light* if the number of matrices $U \in \mathcal{U}$ having 1 on position (r, c) is at most $|\mathcal{U}|/n$. Let the *emptiness* $z_{\mathcal{U}}$ of $\mathcal{U}$ be the number of columns of $A_{\mathcal{U}}$ with at most one 1-entry.

Observation 4.9. *Let $\mathcal{U}$ be a reverse-free set of $n \times k$ word matrices such that $A_{\mathcal{U}}$ has a light 1-entry. Then the set $\mathcal{U}'$ of word matrices of $\mathcal{U}$ not containing the light 1-entry satisfies*

$$
\begin{aligned}
|\mathcal{U}'| &\geq \left(1 - \frac{1}{n}\right) |\mathcal{U}|, \\
|A_{\mathcal{U}'}| &\leq |A_{\mathcal{U}}| - 1 \quad \text{and} \\
z_{\mathcal{U}'} &\geq z_{\mathcal{U}}.
\end{aligned}
$$

Let n_0 be a constant such that for every $n \geq n_0$, $k \leq n$ and $m \geq 5$, every matrix A with $mnk^{1/2}$ 1's contains $n^2m^4/5$ occurrences of S. The existence of n_0 follows from Lemma 4.8.

Claim 4.10. *Let $n \geq n_0$ and let $k \leq n$. Let $\mathcal{U}$ be a reverse-free set of $n \times k$ word matrices with $m_{\mathcal{U}} \geq 5$ and such that $A_{\mathcal{U}}$ has no light 1-entry. Then there exists a set $\mathcal{U}' \subset \mathcal{U}$ satisfying*

$$
\begin{aligned}
|\mathcal{U}'| &\geq \frac{|\mathcal{U}|}{n}, \\
|A_{\mathcal{U}'}| &\leq |A_{\mathcal{U}}| - \frac{2nm_{\mathcal{U}}^3}{5k^{1/2}} \quad \text{and} \\
z_{\mathcal{U}'} &\geq z_{\mathcal{U}} + 1.
\end{aligned}
$$

Proof. The overall matrix $A_{\mathcal{U}}$ contains $n^2m_{\mathcal{U}}^4/5$ occurrences of S. So at least $n^2m_{\mathcal{U}}^4/5$ pairs of 1-entries of $A_{\mathcal{U}}$ are avoided. Thus there is a 1-entry of $A_{\mathcal{U}}$ such that the number of avoided pairs containing this 1-entry is at least

$$
\frac{2n^2m_{\mathcal{U}}^4}{5|A_{\mathcal{U}}|} = \frac{2n^2m_{\mathcal{U}}^4}{5nk^{1/2}m_{\mathcal{U}}} = \frac{2nm_{\mathcal{U}}^3}{5k^{1/2}}.
$$

Let (r, c) be the position of this 1-entry. Let $\mathcal{U}'$ be the set of those matrices from $\mathcal{U}$ that have 1 at position (r, c). We consider a position (r', c') such that

$\{(r,c),(r',c')\}$ is an avoided pair. Every matrix $U' \in \mathcal{U}'$ has 0 at position (r',c'). So also $A_{\mathcal{U}'}$ has 0 at position (r',c'). Therefore $|A_{\mathcal{U}'}| \leq |A_{\mathcal{U}}| - 2nm_{\mathcal{U}}^3/(5k^{1/2})$. Because (r,c) is not a light 1-entry, $|\mathcal{U}'| \geq |\mathcal{U}|/n$. Since $\mathcal{U}'$ contains only word matrices, the matrix $A_{\mathcal{U}'}$ contains only one 1-entry in column r. On the other hand, the 1-entry at position (r,c) is contained in at least one occurrence of S in $A_{\mathcal{U}}$, so $A_{\mathcal{U}}$ contains more than one 1-entry in column r. Thus $z_{\mathcal{U}'} \geq z_{\mathcal{U}} + 1$. □

Claim 4.11. *Let $\mathcal{U}$ be a set of $n \times k$ word matrices, where $n \geq k$. If $\mathcal{U}$ is reverse-free, then*

$$|\mathcal{U}| \leq n^k k^{-k/2+O(k/\log k)}$$

Proof. We first consider the case that the density $m_{\mathcal{U}}$ of the overall matrix is smaller than 10. Since the number of $n \times k$ word matrices contained in $A_{\mathcal{U}}$ is maximized when each of its columns has the same number of 1-entries, we obtain

$$|\mathcal{U}| \leq (10nk^{-1/2})^k$$

and the result holds.

Otherwise, we apply the following procedure on $\mathcal{U}$. We proceed in several steps. Let $\mathcal{U}_i \subset \mathcal{U}$ be the set of word matrices before the step i. Let $\mathcal{U}_1 = \mathcal{U}$. If the overall matrix at the beginning of the step i has a light 1-entry, we obtain $\mathcal{U}_{i+1}$ from $\mathcal{U}_i$ by applying Observation 4.9; otherwise by applying Claim 4.10. Let $m_i \overset{def}{=} m_{\mathcal{U}_i}$ and $A_i \overset{def}{=} A_{\mathcal{U}_i}$. *Light steps* are the steps when Observation 4.9 is applied and *heavy steps* are the remaining ones. The steps are further grouped into *phases*. Phase 1 starts with step $p_1 = 1$. For every $j \geq 2$, phase j starts with step p_j chosen as the smallest index such that $m_{p_j} \leq m_{p_{j-1}}/2$. The last phase is the first phase ℓ that decreases the density of the overall matrix below 10. So at the beginning of the last phase, we have

$$m_{p_\ell} \geq 10.$$

Because each light step decreases the number of 1's in the overall matrix by 1, only at most nk light steps are done during the whole procedure.

It remains to count the heavy steps. At the beginning of a heavy step i of phase j, we have

$$m_i \geq m_{p_j}/2.$$

By Claim 4.10, the heavy step decreases the number of 1-entries in the overall matrix by

$$|A_i| - |A_{i+1}| \geq \frac{2nm_i^3}{5k^{1/2}} \geq \frac{nm_{p_j}^3}{20k^{1/2}}.$$

Since the phase j ends at the moment when at least $|A_{p_j}|/2$ 1-entries are removed, the number of heavy steps of phase j is at most

$$\left\lceil \frac{nk^{1/2}m_{p_j}/2}{nm_{p_j}^3/(20k^{1/2})} \right\rceil = \left\lceil \frac{10k}{m_{p_j}^2} \right\rceil.$$

Each phase shrinks the weight of the overall matrix by a factor of at least 2, so $m_{p_j} \geq m_{p_\ell} 2^{\ell-j}$ for every $j \in \{1,\ldots,\ell\}$. We also have for every such j

$$10 \leq m_{p_j} \leq k^{1/2}.$$

Let t be the total number of heavy steps. We have

$$t \leq \sum_{j=1}^{\ell} \left\lceil \frac{10k}{m_{p_j}^2} \right\rceil \leq \sum_{j=1}^{\ell} \frac{11k}{(m_{p_\ell} 2^{\ell-j})^2} \leq \frac{11k}{m_{p_\ell}^2} \sum_{j=0}^{\infty} 2^{-2j} \leq \frac{11k}{m_{p_\ell}^2} \cdot \frac{4}{3} \leq \frac{k}{6}.$$

Let $\mathcal{U}'$ be the set of word matrices after phase ℓ. During the whole procedure, at most nk light steps and $t \leq k/6$ heavy steps were made. We have

$$|\mathcal{U}'| \geq |\mathcal{U}| \left(1 - \frac{1}{n}\right)^{nk} \left(\frac{1}{n}\right)^{t} \geq |\mathcal{U}| \frac{1}{e^{2k}} n^{-t}. \tag{4.1}$$

The overall matrix $A_{\mathcal{U}'}$ has at most $10nk^{1/2}$ 1-entries and at least t columns with a single 1-entry. The number of $n \times k$ word matrices contained in $A_{\mathcal{U}'}$ is maximized when each of its columns with at least 2 1-entries has the same number of 1-entries. Thus,

$$|\mathcal{U}'| \leq \left(\frac{10nk^{1/2}}{k-t}\right)^{k-t} \leq n^{k-t} \left(\frac{12}{k^{1/2}}\right)^{k} \quad \text{since } t \leq k/6. \tag{4.2}$$

By combining Equations (4.1) and (4.2), we conclude that

$$|\mathcal{U}| \leq n^{k-t} (12k^{-1/2})^k e^{2k} n^t \leq n^k k^{-k/2 + O(k/\log k)}.$$

□

Bibliography

[AER+06] M.H. Albert, M. Elder, A. Rechnitzer, P. Westcott, and M. Zabrocki. On the Stanley–Wilf limit of 4231-avoiding permutations and a conjecture of Arratia. *Adv. Appl. Math.*, 36(2):96–105, 2006.

[AF00] Noga Alon and Ehud Friedgut. On the number of permutations avoiding a given pattern. *Journal of Combinatorial Theory, Series A*, 89(1):133–140, 2000.

[Arr99] Richard Arratia. On the Stanley–Wilf conjecture for the number of permutations avoiding a given pattern. *Electron. J. Comb.*, 6(1):Notes N1, 4 p., 1999.

[ASS89] Pankaj Kumar Agarwal, Micha Sharir, and Peter Shor. Sharp upper and lower bounds for the length of general Davenport–Schinzel sequences. *Journal of Combinatorial Theory, Series A*, 52:228–274, 1989.

[BG91] Daniel Bienstock and Ervin Györi. An extremal problem on sparse 0-1 matrices. *SIAM Journal on Discrete Mathematics*, 4(1):17–27, 1991.

[BHP01] R. C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes, II. *Proceedings of The London Mathematical Society*, 83:532–562, 2001.

[Bón97] Miklós Bóna. Exact enumeration of 1342-avoiding permutations: A close link with labeled trees and planar maps. *J. Comb. Theory, Ser. A*, 80(2):257–272, 1997.

[Bón07] Miklós Bóna. New records in Stanley–Wilf limits. *Eur. J. Comb.*, 28(1):75–85, 2007.

[Bón12] Miklós Bóna. A new upper bound for 1324-avoiding permutations. arXiv:1209.2404 [math.CO], 2012.

[BW00] Eric Babson and Julian West. The permutations $123p_4 \dots p_m$ and $321p_4 \dots p_m$ are Wilf-equivalent. *Graphs Comb.*, 16(4):373–380, 2000.

[BWX07] Jörgen Backelin, Julian West, and Guoce Xin. Wilf-equivalence for singleton classes. *Adv. In Appl. Math.*, 38:133–149, 2007.

[CGN13] Otfried Cheong, Xavier Goaoc, and Cyril Nicaud. Set systems and families of permutations with small traces. *European J. Combin.*, 34(2):229–239, 2013.

[Cib09] Josef Cibulka. On constants in the Füredi–Hajnal and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory, Series A*, 116(2):290–302, 2009.

[CJS12] Anders Claesson, Vít Jelínek, and Einar Steingrímsson. Upper bounds for the Stanley-Wilf limit of 1324 and other layered patterns. *J. Comb. Theory, Ser. A*, 119(8):1680–1691, 2012.

[CK12] Josef Cibulka and Jan Kynčl. Tight bounds on the maximum size of a set of permutations with bounded VC-dimension. *J. Comb. Theory, Ser. A*, 119(7):1461–1478, 2012.

[dM07] Anna de Mier. k-noncrossing and k-nonnesting graphs and fillings of Ferrers diagrams. *Combinatorica*, 27(6):699–720, 2007.

[DS65] Harold Davenport and Andrzej Schinzel. A combinatorial problem connected with differential equations. *American Journal of Mathematics*, 87(3):684–694, 1965.

[Ego81] G. P. Egorychev. Proof of the van der Waerden conjecture for permanents. *Siberian Mathematical Journal*, 22:854–859, 1981.

[EL00] Kimmo Eriksson and Svante Linusson. A combinatorial theory of higher-dimensional permutation arrays. *Adv. in Appl. Math.*, 25(2):194–211, 2000.

[EM64] P. Erdős and J. M. Moon. On the subgraphs of the complete bipartite graph. *Canad. Math. Bull.*, 7:35–39, 1964.

[ES69] P. Erdős and M. Simonovits. Some extremal problems in graph theory. *Col: Math. Soc. J. Bolyai*, 4:377–390, 1969.

[ES96] Alon Efrat and Micha Sharir. A near-linear algorithm for the planar segment-center problem. *Discrete and Computational Geometry*, 16(3):239–257, 1996.

[EV05] M. Elder and V. Vatter. Problems and conjectures presented at the Third International Conference on Permutation Patterns, University of Florida, March 7–11, 2005. arXiv:0505504 [math.CO], 2005.

[Fal81] D. I. Falikman. Proof of the van der Waerden conjecture regarding the permanent of a doubly stochastic matrix. *Mathematical Notes*, 29:475–479, 1981.

[FH92] Zoltán Füredi and Péter Hajnal. Davenport–Schinzel theory of matrices. *Discrete Mathematics*, 103(3):233–251, 1992.

[FKMS10] Zoltán Füredi, Ida Kantor, Angelo Monti, and Blerina Sinaimeri. On reverse-free codes and permutations. *SIAM J. Discrete Math.*, 24(3):964–978, 2010.

[FPS13] J. Fox, J. Pach, and A. Suk. The number of edges in k-quasi-planar graphs. *to appear in SIAM Journal on Discrete Mathematics*, 2013+.

[Ful09] Radoslav Fulek. Linear bound on extremal functions of some forbidden patterns in 0-1 matrices. *Discrete Math.*, 309(6):1736–1739, 2009.

[Für96] Zoltán Füredi. On the number of edges of quadrilateral-free graphs. *J. Comb. Theory, Ser. B*, 68(1):1–6, 1996.

[Gen09] Jesse T. Geneson. Extremal functions of forbidden double permutation matrices. *Journal of Combinatorial Theory, Series A*, 116(7):1235–1244, OCT 2009.

[Ges90] Ira M. Gessel. Symmetric functions and P-recursiveness. *J. Comb. Theory, Ser. A*, 53(2):257–285, 1990.

[HS86] Sergiu Hart and Micha Sharir. Nonlinearity of Davenport–Schinzel sequences and of generalized path compression schemes. *Combinatorica*, 6(2):151–178, 1986.

[Kes09] Balázs Keszegh. On linear forbidden submatrices. *J. Comb. Theory, Ser. A*, 116(1):232–241, 2009.

[KK03] Tomáš Kaiser and Martin Klazar. On growth rates of closed permutation classes. *Electr. J. Combinatorics*, 9(2), 2003.

[Kla92] M. Klazar. A general upper bound in extremal theory of sequences. *Commentationes Mathematicae Universitatis Carolinae*, 33(4):737–746, 1992.

[Kla99] M. Klazar. On the maximum lengths of Davenport–Schinzel sequences. In *Contemporary Trends in Discrete Mathematics, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, volume 49, pages 169–178. AMS, Providence, RI, 1999.

[Kla00] Martin Klazar. The Füredi–Hajnal Conjecture Implies the Stanley–Wilf Conjecture. In *Formal Power Series and Algebraic Combinatorics, Moscow 2000*, pages 250–255. Springer, 2000.

[KM07] Martin Klazar and Adam Marcus. Extensions of the linear bound in the Füredi-Hajnal conjecture. *Adv. Appl. Math.*, 38(2):258–266, 2007.

[Knu68] D.E. Knuth. *The art of computer programming. Vol. 1: Fundamental algorithms.* Addison–Wesley Series in Computer Science and Information Processing. London: Addison–Wesley Publishing Company. XXII, 634 p. , 1968.

[KV94] Martin Klazar and Pavel Valtr. Generalized Davenport–Schinzel sequences. *Combinatorica*, 14:463–476, 1994.

[Kyn12] Jan Kynčl. Improved enumeration of simple topological graphs. arXiv:1212.2950 [math.CO], 2012.

[LL11] Nathan Linial and Zur Luria. An upper bound on the number of high-dimensional permutations. arXiv:1106.0649 [math.CO], 2011.

[Mat02] Jiří Matoušek. *Lectures on Discrete Geometry.* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2002.

[ML10] Neal Madras and Hailong Liu. Random pattern-avoiding permutations. Lladser, Manuel E. (ed.) et al., Algorithmic probability and combinatorics. Papers from the AMS special sessions, Chicago, IL, USA, October 5–6, 2007 and Vancouver, BC, Canada, October 4–5, 2008. Providence, RI: American Mathematical Society (AMS). Contemporary Mathematics 520, 173-194 (2010)., 2010.

[MN98] Jiří Matoušek and Jaroslav Nešetřil. *Invitation to Discrete Mathematics*. Oxford University Press, 1998.

[MT04] Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory, Series A*, 107(1):153–160, 2004.

[MW99] B. D. McKay and I. M. Wanless. Most Latin squares have many subsquares. *J. Combin. Theory Ser. A*, 86(2):322–347, 1999.

[Niv10] Gabriel Nivasch. Improved bounds and new techniques for Davenport–Schinzel sequences and their generalizations. *Journal of the Association for Computing Machinery*, 57(3):1–44, 2010.

[Pet11a] S. Pettie. Degrees of Nonlinearity in Forbidden 0-1 Matrix Problems. *Discrete Mathematics*, 311:2396–2410, 2011.

[Pet11b] S. Pettie. Generalized Davenport-Schinzel Sequences and Their 0-1 Matrix Counterparts. *Journal of Combinatorial Theory, Series A*, 118(6):1863–1895, 2011.

[Pet11c] S. Pettie. On the structure and composition of forbidden sequences, with geometric applications. In *Symposium on Computational Geometry*, pages 370–379, 2011.

[Pet12] S. Pettie. Tightish Bounds on Davenport–Schinzel Sequences. arXiv:1204.1086v1 [cs.DM], 2012.

[PS91] János Pach and Micha Sharir. On vertical visibility in arrangements of segments and the queue size in the Bentley–Ottmann line sweeping algorithm. *SIAM J. Comput.*, 20(3):460–470, 1991.

[PT06] János Pach and Gábor Tardos. Forbidden paths and cycles in ordered graphs and matrices. *Israel Journal of Mathematics*, 155:359–380, 2006.

[Raz00] Ran Raz. VC-Dimension of Sets of Permutations. *Combinatorica*, 20(2):241–255, 2000.

[SA95] Micha Sharir and Pankaj Kumar Agarwal. *Davenport–Schinzel Sequences and Their Geometric Applications*. Cambridge University Press, Cambridge, MA, 1995.

[Sid91] A. Sidorenko. Inequalities for functionals generated by bipartite graphs. *Diskret. Mat.*, 3:50–65, 1991.

[SS85] Rodica Simion and Frank W. Schmidt. Restricted permutations. *Eur. J. Comb.*, 6:383–406, 1985.

[Sta96] Zvezdelina Stankova. Classification of forbidden subsequences of length 4. *Eur. J. Comb.*, 17(5):501–517, 1996.

[Ste12] Einar Steingrímsson. Some open problems on permutation patterns. arXiv:1210.7320 [math.CO], 2012.

[SW02] Zvezdelina Stankova and Julian West. A new class of Wilf-equivalent permutations. *J. Algebr. Comb.*, 15(3):271–290, 2002.

[Tar05] Gábor Tardos. On 0-1 matrices and small excluded submatrices. *Journal of Combinatorial Theory, Series A*, 111(2):266–288, 2005.

[Val97] Pavel Valtr. Graph drawing with no k pairwise crossing edges. In *Graph drawing (Rome, 1997)*, volume 1353 of *Lecture Notes in Comput. Sci.*, pages 205–218. Springer, Berlin, 1997.

[vLW92] J. H. van Lint and R. M. Wilson. *A course in combinatorics.* Cambridge University Press, Cambridge, 1992.

[Wes90] Julian West. *Permutations with forbidden subsequences and stack-sortable permutations.* PhD thesis, M.I.T., 1990.