Large homogeneous submatrices

Dániel Korándi *   János Pach †‡   István Tomon §‡

Abstract

A matrix is homogeneous if all of its entries are equal. Let P be a 2×2 zero-one matrix that is not homogeneous. We prove that if an $n \times n$ zero-one matrix A does not contain P as a submatrix, then A has an $cn \times cn$ homogeneous submatrix for a suitable constant $c > 0$. We further provide an almost complete characterization of the matrices P (missing only finitely many cases) such that forbidding P in A guarantees an $n^{1-o(1)} \times n^{1-o(1)}$ homogeneous submatrix. We apply our results to chordal bipartite graphs, totally balanced matrices, halfplane-arrangements and string graphs.

1 Introduction

Zero-one matrices play an important role in discrete mathematics, as they can be used to represent (bipartite) graphs, hypergraphs, systems of incidences, and many other binary relations. In such settings, the circumstances often force structural restrictions. In this paper, we analyze the structure of matrices that do not contain a given submatrix P, and show that forbidding P often forces a large all-0 or all-1 submatrix. With a slight abuse of notation, the letter c appearing in different statements stands for unrelated positive constants.

A matrix is *homogeneous* if all of its entries are equal, and *inhomogeneous* otherwise. We will also say that a matrix A *contains* another matrix P if P is a submatrix of A, and that A is *P-free* if P is not a submatrix of A. Our first result shows that if P is a 2×2 matrix whose entries are not all 0 or all 1, then every P-free zero-one matrix contains a linear-size homogeneous submatrix.

Theorem 1.1. *Let P be an inhomogeneous 2×2 zero-one matrix. Then every P-free $n \times n$ zero-one matrix A contains a homogeneous $cn \times cn$ submatrix, for a suitable constant $c > 0$.*

As we will see below, this result does not hold when P is the all-0 or the all-1 2×2 matrix. We can, however, extend Theorem 1.1 to $2 \times k$ matrices by making a small sacrifice on the size of the homogeneous submatrix.

Theorem 1.2. *Let P be a $2 \times k$ zero-one matrix that does not contain a 2×2 homogeneous submatrix. Then every P-free $n \times n$ zero-one matrix A contains a homogeneous $n^{1-o(1)} \times cn$ submatrix, for a suitable constant $c > 0$.*

Of course, one can obtain an analogous result for $k \times 2$ matrices by working with the transposes. In particular, we find a homogeneous $n^{1-o(1)} \times n^{1-o(1)}$ submatrix for any $2 \times k$ or $k \times 2$ matrix P with no 2×2 homogeneous submatrix. Moreover, every $1 \times k$ matrix can be extended to such a $2 \times k$ matrix, so this also holds for $1 \times k$ and $k \times 1$ matrices.

We should also point out that, as permuting rows or columns does not affect homogeneous submatrices, the same results hold if we only assume that A can be made P-free by reordering its rows and columns.

Our problem arises naturally in numerous combinatorial and geometrical settings. When A represents a bipartite graph, P corresponds to a forbidden induced (ordered) subgraph, and a homogeneous submatrix is a complete or empty bipartite subgraph. When A represents an incidence relation in geometry, P is often a geometrically impossible pattern, and a homogeneous submatrix corresponds to two completely intersecting or disjoint families. When A is the incidence matrix of a hypergraph, a homogeneous submatrix gives a set of hyperedges and a completely disjoint or completely contained set of vertices. We list a few specific applications to chordal bipartite graphs, totally balanced matrices, halfplane-arrangements and string graphs in Section 9.

Several closely related problems have been studied in the literature, including certain Erdős-Hajnal type questions and the Turán problem for ordered graphs and forbidden patterns. We discuss some connections and differences in Section 10.

As mentioned above, it is not true that forbidding any submatrix P forces an almost linear-size homogeneous submatrix.

Definition 1.3. *A zero-one matrix P is called* acyclic *if every submatrix of P has a row or column containing at most one 1-entry. The* complement *of P is the matrix P^c obtained from P by replacing the 1-entries with 0s and the 0-entries with 1s. We say that P is* simple *if both P and P^c are acyclic.*

It is easy to see that P is acyclic if and only if the bipartite graph with biadjacency matrix P is acyclic. If P is not simple, then there are P-free zero-one matrices with only small homogeneous submatrices. The (fairly standard) probabilistic construction will be given in Section 3.

Proposition 1.4. *Let P be a zero-one matrix. If P is not simple, then there is a P-free $n \times n$ zero-one matrix A with no homogeneous $n^{1-\varepsilon} \times n^{1-\varepsilon}$ submatrix for every large enough n, where $\varepsilon = \varepsilon(P)$ is a positive constant.*

Proposition 1.4 shows that Theorems 1.1 and 1.2 are optimal in terms of the matrices covered: the remaining 2×2 or $2 \times k$ matrices are not simple, so these statements cannot hold for them. In fact, Theorem 1.2 and Proposition 1.4 almost completely characterize which forbidden matrices force an almost-linear homogeneous submatrix, because they only miss a finite number of simple matrices. Indeed, a simple $k \times \ell$ matrix can contain at most $k + \ell - 1$ 0-entries and at most $k + \ell - 1$ 1-entries, so it must satisfy $2k + 2\ell - 2 \geq k\ell$, or, equivalently, $(k-2)(\ell-2) \leq 2$. So, apart from the matrices treated in Theorem 1.2, only 3×3, 3×4 and 4×3 matrices can be simple.

We believe that a similar statement should hold for the remaining simple matrices, as well. In fact, we make the following stronger conjecture.

Conjecture 1.5. *Let P be a simple zero-one matrix. Then every P-free $n \times n$ zero-one matrix contains an $cn \times cn$ homogeneous submatrix, for a suitable constant $c > 0$.*

Much of the difficulty in our results comes from the ordered structure of matrices. We can obtain better results if we relax the notion of matrices to "unordered" matrices, where the order of the rows and the columns does not matter. We can then say that a zero-one matrix is *unordered P-free* if it does not contain any submatrix whose rows and columns can be permuted to obtain P. We show that Theorem 1.2 holds for unordered $2 \times k$ matrices.

Theorem 1.6. *Let P be a $2 \times k$ zero-one matrix that does not contain a 2×2 homogeneous submatrix. Then every unordered P-free $n \times n$ zero-one matrix A contains a homogeneous $cn \times cn$ submatrix, for a suitable constant $c > 0$.*

Results about unordered matrices can be thought of as results about bipartite graphs. In the language of graphs, Theorem 1.6 implies the following statement: let $H_{s,t}$ be a star of size s and a star of size t glued together at one of their leaves, and let $H_{s,t}^*$ be the union of $H_{s,t}$ and an isolated vertex. Let H be an induced subgraph of $H_{s,t}^*$, and let $G = (A \cup B, E)$ be an induced H-free bipartite graph with $|A| = |B| = n$. Then there are linear-size subsets $A_0 \subseteq A$ and $B_0 \subseteq B$ such that $A_0 \cup B_0$ induces either a complete or an empty bipartite graph in G. This latter statement has been proved independently by Axenovich, Tompkins, and Weber [2].

Our paper is organized as follows. In Section 2, we state a number of further results, which imply Theorems 1.1 and 1.2, but might be of interest on their own. The proof of these (positive) results are given in Sections 4 to 7. Our negative result, Proposition 1.4, is proved in Section 3. Finally, we prove Theorem 1.6 in Section 8.

We finish the paper with a few applications in Section 9 and some further connections and remarks in Section 10.

2   Forbidden submatrices

Our first result is about $2 \times k$ matrices that contain a 0-entry and a 1-entry in every column, establishing Theorems 1.1 and 1.2 in this special case.

Theorem 2.1. *Let P be a $2 \times k$ zero-one matrix without any homogeneous column. Then every P-free $n \times n$ zero-one matrix contains a $c\frac{n}{k} \times c\frac{n}{k}$ homogeneous submatrix, for a suitable constant $c \geq 10^{-6}$.*

As rotation and taking complements does not affect the problem, there is essentially one simple 2×2 matrix not covered by this theorem: $Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. When P cannot be rotated into a $2 \times k$ matrix without homogeneous columns, the problem becomes more difficult. Q is the only such matrix where we can prove a linear lower bound.

Theorem 2.2. *Let $Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Then every Q-free $n \times n$ zero-one matrix contains an $cn \times cn$ homogeneous submatrix, for a suitable constant $c \geq 1/20$.*

In the general case, we can show the following, somewhat weaker result.

Theorem 2.3. *Let P be a simple $2 \times k$ zero-one matrix. Then every P-free $n \times n$ zero-one matrix contains an $n^{1-o(1)} \times cn$ homogeneous submatrix, for a suitable constant $c > 0$.*

Theorem 1.1 then follows from Theorems 2.1 and 2.2, and Theorem 1.2 is equivalent to Theorem 2.3. Theorems 2.1, 2.2 and 2.3 are proved in Sections 4, 5, and 6, respectively.

Note that Conjecture 1.5 is invariant under taking complements: the complement of a P-free zero-one matrix A is P^c-free, and if P is simple, then so is P^c. We may therefore assume that 0 is the majority entry in A, and then try to find a large all-0 submatrix in it. Indeed, this is the approach we take to prove Theorems 2.1 to 2.3. More generally, we believe that the following strengthening of Conjecture 1.5 might also be true.

3

Conjecture 2.4. *Let P be an acyclic zero-one matrix. Then for every $\varepsilon > 0$, there is a $\delta > 0$, such that every P-free $n \times n$ zero-one matrix with at least εn^2 0-entries contains a $\delta n \times \delta n$ all-0 submatrix.*

Another immediate corollary of this conjecture would be the following:

Conjecture 2.5. *Let P be an acyclic zero-one matrix. Then every $n \times n$ zero-one matrix that is both P-free and P^c-free contains an $cn \times cn$ homogeneous submatrix, for a suitable constant $c > 0$.*

We can prove these conjectures in the special case when P has no column with more than one 1-entries.

Theorem 2.6. *Let P be a zero-one matrix such that every column of P has at most one 1-entry. Then every $n \times n$ zero-one matrix that is both P-free and P^c-free contains a $cn \times cn$ homogeneous submatrix, for some $c > 0$.*

Note that Theorem 2.1 can also be obtained, with slightly weaker constants, as a corollary of this result (by applying Theorem 2.6 to the concatenation of P and P^c). The proof can be found in Section 7.

3 Notation, preliminaries–Proof of Proposition 1.4

Throughout this paper, we use the following notation. When A is a matrix, $A(i, j)$ denotes the entry in the i'th row and j'th column. Sometimes we make no distinction between rows and their indices, and refer to the i'th row as "row i" (and, in a similar manner, for columns). We denote the submatrix in the intersection of rows X and columns Y (the submatrix induced by X and Y) by $A[X \times Y]$.

We use two natural correspondences between zero-one matrices and graphs. The *biadjacency matrix* of a bipartite graph $G = (A \cup B, E)$ is the zero-one matrix whose rows are indexed by A, columns are indexed by B, and the (a, b) entry is 1 for $a \in A$ and $b \in B$ if and only if $ab \in E$. The *incidence matrix* of a graph $G = (V, E)$ is the zero-one matrix whose rows are indexed by V, columns are indexed by E, and the (v, e) entry is 1 if and only if e is incident to v.

For two subsets $X, Y \subseteq [n]$, we write $X < Y$ to denote that $x < y$ for every $x \in X, y \in Y$. When $Y = \{y\}$, we may simply write $X < y$. We systematically omit floor and ceiling signs whenever they are not essential.

We start by proving that only acyclic forbidden matrices can force large all-0 submatrices. This also shows that Conjecture 2.4 can only hold for acyclic P.

Proposition 3.1. *Let P be a $k \times \ell$ zero-one matrix. If P is not acyclic, then there is a P-free $n \times n$ zero-one matrix A with at least $n^2/2$ 0-entries, but no homogeneous $n^{1-\varepsilon} \times n^{1-\varepsilon}$ submatrices for every large enough n, where $\varepsilon = \varepsilon(P)$ is a positive constant.*

Proof. We may assume that every row and column of P contains at least two 1-entries, as otherwise we can replace P with a submatrix. In particular, we have $k, \ell \geq 2$, and P contains at least $k + \ell$ 1-entries.

Let A_0 be a random $n \times 2n$ matrix, where each entry is independently set to 1 with probability $p = \frac{1}{4}n^{-1+\frac{1}{k+\ell}}$, and set to 0 otherwise. First of all, note that the expected number of 1-entries is $2n^2 p = \frac{1}{2}n^{1+\frac{1}{k+\ell}} < n^2/8$ if n is large enough, so the probability that A_0 has more than $n^2/2$ 1-entries is at most $1/4$. Also, the expected number of submatrices identical to P in A_0 is at

most $\binom{n}{k}\binom{2n}{\ell}p^{k+\ell} < (2np)^{k+\ell} < n/4$. So with probability at least $1/4$, there are at most n such submatrices. Finally, the probability that A_0 contains a homogeneous $m \times m$ matrix is at most

$$\binom{n}{m}\binom{2n}{m}\left(p^{m^2} + (1-p)^{m^2}\right) \leq (2n)^{2m}(1-p)^{m^2} \leq e^{4m\log n - pm^2} < 1/4$$

if $4\log n - pm < -2$, which holds for $m = n^{1-\varepsilon}$ whenever $\varepsilon < \frac{1}{k+\ell}$ and n is large enough. So there is an $n \times 2n$ matrix that contains at most n submatrices identical to P and no homogeneous $m \times m$ submatrix. Then we can delete n columns to obtain the P-free matrix A we were looking for. $\qquad\square$

Proof of Proposition 1.4. Let us apply Proposition 3.1 to P or P^c (whichever is not acyclic) to get A with no homogeneous $n^{1-\varepsilon} \times n^{1-\varepsilon}$ submatrix. Then A or A^c (whichever is P-free) will work. $\qquad\square$

Definition 3.2. *We say that a zero-one matrix P is (ε, δ)-good if for all n, every $n \times n$ P-free matrix with at least εn^2 0-entries contains a $\delta n \times \delta n$ all-0 submatrix.*

By convention, every matrix contains the 0×0 all-0 submatrix, so every P is $(\varepsilon, 0)$-good for every ε. We prove our main results by showing that certain matrices P are (ε, δ)-good for some $\delta > 0$. Let us start with a simple case.

Proposition 3.3. *The all-1 $1 \times k$ matrix $P = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix}$ is $(0, 1/k)$-good.*

Proof. Without assuming anything about the density, we can find an $\frac{n}{k} \times \frac{n}{k}$ all-0 matrix in any $\frac{n}{k}$ rows of an $n \times n$ P-free matrix. Indeed, as every row contains at most $k - 1$ 1-entries, any $\frac{n}{k}$ rows induce at least $n - \frac{n(k-1)}{k} = \frac{n}{k}$ columns with only 0-entries. $\qquad\square$

Of course, if a matrix is (ε, δ)-good, then it is also (ε', δ)-good for any $\varepsilon' \geq \varepsilon$. The next lemma shows that adding an all-0 row or column at a border of a matrix does not change goodness.

Lemma 3.4. *Let P be a $k \times \ell$ zero-one matrix, and let P' be the $k \times (\ell + 1)$ matrix obtained from P by appending a new last column of 0-entries. If P is (ε, δ)-good for some $\varepsilon \geq 0$, then P' is $(2\varepsilon, \delta\varepsilon)$-good.*

Proof. Let A be a P'-free $n \times n$ matrix with at least $2\varepsilon n^2$ 0-entries. We will find a dense submatrix with an all-0 last column, and then apply the goodness property of P to get the large homogeneous submatrix.

Define A' as the matrix obtained from A by replacing the first εn 0-entries of each row by 1-entries (if a row has fewer than εn 0-entries, then it becomes a row with all 1's). Then A' has at least εn^2 0-entries, so it must contain a column with at least εn 0-entries. If column j is such a column, let I be a set of εn rows with a 0-entry in the j'th column, and let J_0 be the first $j - 1$ columns. By the definition of A', every row of $B_0 = A[I \times J_0]$ has at least εn 0-entries, so in total, B_0 contains at least $\varepsilon^2 n^2$ 0-entries. Now let $J \subseteq J_0$ be the εn columns with the most 0-entries in them. Then $B = A[I \times J]$ is an $\varepsilon n \times \varepsilon n$ matrix with at least $\varepsilon^3 n^2$ 0-entries.

Note that B is P-free, since we could otherwise add 0's in the j'th column to get a copy of P' in A. As P is (ε, δ)-good, B must contain a $\delta\varepsilon n \times \delta\varepsilon n$ all-0 submatrix. $\qquad\square$

4 Matrices with no homogeneous columns–Proof of Theorem 2.1

In this section, we prove Theorem 2.1 by showing that every $2 \times k$ matrix with no homogeneous columns satisfies Conjecture 2.4. We first prove this for a special class of "checkerboard" matrices. Let P_k denote the $2 \times k$ matrix defined by $P_k(i, j) = 1$ if $i + j$ is even, and $P_k(i, j) = 0$ otherwise.

The main concern of this section is to establish that for every $\varepsilon > 0$, there is a $\delta > 0$ such that P_{2k} is (ε, δ)-good. The general case will follow easily by observing that every $2 \times k$ matrix with no homogeneous columns is a submatrix of P_{2k}.

Note that P_{2k} is the concatenation of k copies of $P_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. We first consider P_2-free families.

Lemma 4.1. *Let $\varepsilon > 0$, and suppose that A is an $n \times n$ zero-one matrix with at least εn^2 0-entries. Then at least one of the following statements holds.*

1. *A contains an $\frac{\varepsilon n}{8} \times \frac{\varepsilon n}{8}$ all-0 submatrix.*

2. *At least $\frac{\varepsilon^2 n^2}{64}$ different pairs of rows of A contain P_2 as a submatrix.*

Proof. Let $t = \frac{\varepsilon n}{8}$. First, we find a $2t \times 2t$ submatrix of A such that its first row and column contain only 0's, moreover, each of these 0-entries is preceded by $2t$ other 0-entries in their rows and columns in A.

Let A' be the matrix obtained from A by replacing the first $2t$ 0-entries of each row and column with 1-entries. As at most $4tn$ 0-entries are lost, A' still has at least $\frac{\varepsilon n^2}{2}$ 0-entries. Now let A'' be the matrix obtained from A' by replacing the last $2t - 1$ 0-entries of each row and column with 1-entries. By the same argument, A'' has at least $2n$ 0-entries.

Take a 0-entry in A'', say in the i_1'th row and j_1'th column. By the definition of A'', we must have a set $J > j_1$ of $2t - 1$ columns such that the i_1'th row of A' contains a 0 in these columns, and similarly, there we must have a set $I > i_1$ of $2t - 1$ rows such that the j_1'th column of A' contains a 0 in these rows. So, the submatrix $A'[(\{i_1\} \cup I) \times (\{j_1\} \cup J)]$ is all-0 in its first row and column. Also, by the definition of A', each row $i \in I$ contains $2t$ 0-entries in A in some columns Y_i preceding the columns of J, and similarly, each column $j \in J$ contains $2t$ 0-entries in some rows $X_j < I$.

If $A[I \times J]$ has t rows without a 1-entry, then it contains a $t \times t$ all-0 submatrix, establishing 1. Hence, we may assume that at least t rows in $A[I \times J]$ contain a 1-entry.

Let $i \in I, j \in J$ be such that $A(i,j) = 1$, and look at the $2t \times 2t$ submatrix $A[X_i \times Y_j]$. Again, if this has t rows without a 1-entry, then A contains a $t \times t$ all-0 submatrix, and we are done. Otherwise, there are 1-entries in t different rows of $A[X_i \times Y_j]$. However, if for some $x \in X_i, y \in Y_i$, the entry $A(x,y)$ is 1, then $A[\{x,i\} \times \{y,j\}] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. For any choice of (i,j), there is such an (x,y) in t different rows, so we find P_2 in at least t^2 different row pairs, establishing 2. □

Lemma 4.2. *For every $\varepsilon > 0$, P_{2k} is $(\varepsilon, \frac{\varepsilon^4}{10^4 k})$-good.*

Proof. Suppose A is a P_{2k}-free $n \times n$ zero-one matrix. Let $s = 400k/\varepsilon^3$, and divide A into $\frac{n}{s} \times \frac{n}{s}$ blocks $A_{i,j} = A[I_i \times I_j]$, where I_k is the interval $[\frac{(k-1)n}{s} + 1, \frac{kn}{s}]$, for every $i, j, k \in [s]$. We say that $(i,j) \in [s]^2$ is *heavy* if $A_{i,j}$ contains at least $\frac{\varepsilon n^2}{2s^2}$ 0-entries. If N denotes the number of heavy pairs, then we can bound the number of 0-entries in A as follows:

$$\varepsilon n^2 \leq N \cdot \frac{n^2}{s^2} + s^2 \cdot \frac{\varepsilon n^2}{2s^2}.$$

Consequently, $N \geq \varepsilon s^2/2$.

This means that for some $i_0 \in [s]$, there is a set $J \subseteq [s]$ of at least $t = \varepsilon s/2$ indices such that (i_0, j) is heavy for every $j \in J$. Let R_j be the set of pairs $\{r, q\} \in [n/s]^{(2)}$ such that rows r and q in $A_{i_0,j}$ together contain P_2. If (i_0, j) is heavy, then by Lemma 4.1 (applied with parameters $\varepsilon/2$

6

and n/s), either $|R_j| \geq \frac{(\varepsilon n/2s)^2}{64}$, or $A_{i_0,j}$ contains an $\frac{\varepsilon n}{16s} \times \frac{\varepsilon n}{16s}$ all-0 submatrix. In the latter case, we are done, so we may assume the former holds for every $j \in J$. Now

$$t \cdot \frac{(\varepsilon n/2s)^2}{64} = \frac{\varepsilon^3 s}{256} \cdot \frac{(n/s)^2}{2} > k \binom{n/s}{2}$$

implies that some pair $\{r, q\}$ is contained in at least k of the sets R_j, say in $R_{j_1}, \ldots, R_{j_k}$. Then P_{2k} is a submatrix of the union of the matrices $A_{i_0,j_1}, \ldots, A_{i_0,j_k}$ in the rows indexed by r and q, which is a contradiction. $\qquad \square$

Proof of Theorem 2.1. Every $2 \times k$ matrix P with no homogeneous columns is contained in P_{2k}, so if a matrix is P-free, then it is also P_{2k}-free.[1] Similarly, every P^c-free matrix is P_{2k}-free, because P^c also has no homogeneous columns.

If A is P-free, then A^c is P^c-free, so both A and A^c are P_{2k}-free. One of A and A^c will contain at least $n^2/2$ 0-entries, so we can apply Lemma 4.2 with $\varepsilon = 1/2$ to find an $\frac{n}{20^4 k} \times \frac{n}{20^4 k}$ homogeneous submatrix in A. $\qquad \square$

Let $f_k(\varepsilon) = \sup\{\delta : P_{2k} \text{ is } (\varepsilon, \delta)\text{-good}\}$, that is, $f_k(\varepsilon)$ is the largest δ such that for every n, every $n \times n$ P_{2k}-free matrix with εn^2 0-entries contains a $\delta n \times \delta n$ all-0 matrix. One might wonder what the order of $f_k(\varepsilon)$ is. Lemma 4.1 shows that $f_1(\varepsilon) = \Theta(\varepsilon)$ (the upper bound $f_1(\varepsilon) \leq \varepsilon$ is trivial), while Lemma 4.2 implies $f_k(\varepsilon) = \Omega(\varepsilon^4)$ for $k \geq 2$. It might seem reasonable to conjecture that $f_k(\varepsilon) = \Theta(\varepsilon)$ also holds for $k \geq 2$. However, this is not true, already for $k = 2$: Füredi and Hajnal [18] proved that for every positive integer m, there is an $m \times m$ matrix B such that B does not contain either of $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} * & 0 & * & 0 \\ 0 & * & 0 & * \end{pmatrix}$ as a submatrix (where $*$ can be either 0 or 1), but B contains $\Omega(m\alpha(m))$ 0-entries, where $\alpha(m)$ is the slowly growing inverse Ackermann function. For $\varepsilon = \Omega(\alpha(m)/m)$ and every $n > m$, we can construct the $n \times n$ matrix A by replacing each 1-entry of B with an $\frac{n}{m} \times \frac{n}{m}$ all-1 matrix, and each 0-entry of B with an $\frac{n}{m} \times \frac{n}{m}$ all-0 matrix. Then A is P_4-free, it has at least εn^2 0-entries, but it does not contain any all-0 submatrix with more than $\frac{n}{m}$ rows and columns. As $\frac{1}{m} = O(\frac{\varepsilon}{\alpha(1/\varepsilon)})$, we have $f_2(\varepsilon) = O(\frac{\varepsilon}{\alpha(1/\varepsilon)})$.

It would be interesting to determine the true order of magnitude of $f_2(\varepsilon)$. We believe the answer should be closer to the upper bound $O(\frac{\varepsilon}{\alpha(1/\varepsilon)})$.

5 The 2×2 matrix with one 1 in the corner–Proof of Theorem 2.2

In this section, we establish Theorem 2.2. As before, we achieve this by showing a density result: we prove that both $Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and its complement satisfy Conjecture 2.4.

More generally, let Q_k be the the the $2 \times (k+1)$ matrix such that $Q_k(1, i) = 1$ for $i = 1, \ldots, k$, and all other entries are 0. For example, $Q = Q_1$, and $Q_3 = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Proposition 3.3 and Lemma 3.4 easily imply that Q_k is $(\varepsilon, \varepsilon^2/k)$-good for every ε. In this case, we can actually gain a factor of ε:

Lemma 5.1. Q_k *is* $(\varepsilon, \frac{\varepsilon}{2k})$*-good for every* $\varepsilon \geq 0$.

Proof. Let A be a Q_k-free $n \times n$ matrix with at least εn^2 0-entries, and let A' be the matrix obtained from A by replacing the first $\varepsilon n/2$ 0-entries in each row and column with 1's. It is easy to see that

---

[1]Note that this observation combined with Lemma 4.2 also implies that every such P is $(\varepsilon, \frac{\varepsilon^4}{10^4 k})$-good.

7

fewer than εn^2 entries were changed, so $A'(i_0, j_0) = 0$ for some $i_0, j_0 \in [n]$. By the definition of A', we then have sets $I, J \subseteq [n]$ of size $\varepsilon n/2$ such that $I < i_0$ and $J < j_0$, and for every $i \in I$ and $j \in J$, $A(i, j_0) = A(i_0, j) = 0$.

Now $A[I \times J]$ is an $\frac{\varepsilon n}{2} \times \frac{\varepsilon n}{2}$ matrix, and as A is Q_k-free, it does not contain a $1 \times k$ all-1 submatrix. Then, by Proposition 3.3, it has an $\frac{\varepsilon n}{2k} \times \frac{\varepsilon n}{2k}$ all-0 submatrix. $\square$

The difficult part is to show that for every $\epsilon > 0$, Q^c is also (ε, δ)-good for some $\delta > 0$. We prove this in the next lemma.

Lemma 5.2. *Let A be an $n \times n$ zero-one matrix with at least εn^2 1-entries. If A does not contain Q, then it has an $\frac{\varepsilon n}{18} \times \frac{\varepsilon n}{18}$ all-1 submatrix.*

Proof. For an index $i \in [n]$, let X_i denote the submatrix formed by the first i columns of A and let Y_i denote the submatrix of the last $n - i$ columns. Then for some i, both X_i and Y_i contain at least $\varepsilon n^2/3$ 1-entries. Note that this implies, in particular, that both X_i and Y_i have at least $\varepsilon n/3$ columns. Also, X_i has at least $\varepsilon n/6$ rows containing at least $\varepsilon n/6$ 1-entries. Indeed, otherwise X_i would contain fewer than $\frac{\varepsilon n}{6} \cdot n + (n - \frac{\varepsilon n}{6}) \cdot \frac{\varepsilon n}{6} < \varepsilon n^2/3$ 1-entries in total, which is not the case. Let X be the submatrix of X_i consisting of $\varepsilon n/6$ such rows, and let Y be an $\frac{\varepsilon n}{6} \times \frac{\varepsilon n}{6}$ submatrix of the same rows in Y_i.

Now let us define the graph G on the 0-entries of Y as vertices, where we connect two 0-entries by an edge if they are in the same row or the same column of Y. For a vertex v in G, we define $r(v)$ and $c(v)$ as the row and column of v, respectively. We say that a path $v_1 \ldots v_k$ in G is row-monotone if $r(v_1) \leq \cdots \leq r(v_k)$. This notion is motivated by the following claim.

Claim 5.3. *Let $v \in G$ be a vertex of G, and let $U = \{u_1, \ldots, u_s\}$ be the set of vertices that can be reached from v via a row-monotone path. Then X contains a $t \times \frac{\varepsilon n}{6}$ all-1 submatrix, where $t = |r(U)|$ is the number of different rows of U.*

Proof. Let $u \in U$ be a vertex that can be reached from v via a row-monotone path $u_0 u_1 \ldots u_k$, where $u_0 = v$ and $u_k = u$. We are going to show that if $A(r(v), x) = 1$ for some column x of X, then $A(r(u), x) = 1$, as well. In fact, we will show $A(r(u_i), x) = 1$ for every i, by induction.

Assume this holds for some i (the case $i = 0$ is trivial). If $r(u_i) = r(u_{i+1})$, then there is nothing to prove. Otherwise, $r(u_i) < r(u_{i+1})$ and $c(u_i) = c(u_{i+1})$ by the definition of the path. Let us look at the submatrix $A[\{r(u_i), r(u_{i+1})\} \times \{x, c(u_i)\}]$. The entries in the second column are 0 by the definition of G, and the top left entry is 1 by assumption. But this submatrix cannot be Q, so the bottom left entry $A(r(u_{i+1}), x)$ must also be 1, as needed.

This shows that in X, the rows of $r(U)$ have 1-entries wherever $r(v)$ does. The row $r(v)$, like every row of X, contains at least $\varepsilon n/6$ 1-entries, so the rows of $r(U)$ together produce a $t \times \frac{\varepsilon n}{6}$ all-1 submatrix. $\square$

Claim 5.3 shows that it would be enough to find a vertex in G that sends monotone paths to at least $\varepsilon n/18$ different rows. The next claim shows that each connected component of G has a vertex v that reaches the whole component via monotone paths.

Claim 5.4. *Let C be a connected component of G and let $v \in C$ be a vertex such that $r(v)$ is smallest. Then for every vertex $u \in C$, there is a row-monotone path from v to u.*

Proof. Let $P = v_0 \ldots v_k$ be a v-u walk in C that minimizes $\sum_{w \in P} r(w)$. We will show that P is a row-monotone path. First, we establish the following simple properties for every such minimal path:

8

1. P has no three collinear vertices, i.e., there is no i such that $c(v_{i-1}) = c(v_i) = c(v_{i+1})$ or $r(v_{i-1}) = r(v_i) = r(v_{i+1})$.

2. There is no "bottom right corner" in P, i.e., there is no i such that $r(v_{i-1}) < r(v_i)$ and $c(v_i) > c(v_{i+1})$, and there is no i with $c(v_{i-1}) < c(v_i)$ and $r(v_i) > r(v_{i+1})$.

The first property is clear: we would get a better v-u walk by simply deleting v_i from P. For the second property, suppose there is an i satisfying $r(v_{i-1}) < r(v_i)$ and $c(v_i) > c(v_{i+1})$, and look at the 2×2 submatrix $M = A[\{r(v_{i-1}), r(v_i)\} \times \{c(v_{i+1}), c(v_i)\}]$. Using $c(v_{i-1}) = c(v_i)$ and $r(v_i) = r(v_{i+1})$, we see that P contains all entries of this submatrix, except for the top left entry. All vertices in P are 0-entries, so $A(r(v_{i-1}), c(v_{i+1})) = 0$, as well, for otherwise $M = Q$. Then we could replace v_i with the vertex corresponding to this top left entry, and get a new P with smaller $\sum_{w \in P} r(w)$. The other case of property 2 can be proved analogously.

Now let j be the smallest index such that $r(v_j) \neq r(v)$. By the definition of v, we have $r(v_j) > r(v_{j-1})$. We can show by induction that from v_{j-1} on, P alternately moves downwards and to the right. Indeed, property 1 shows that the path changes direction after each edge. Now suppose that at some point it moves downwards, i.e., $r(v_{i-1}) < r(v_i)$ (as is the case for $i = j$). Then according to property 2, we cannot move towards the left, so we must have $c(v_i) < c(v_{i+1})$. On the other hand, if the path moves to the right, i.e., $c(v_{i-1}) < c(v_i)$, then the second case of property 2 forbids a move upwards in the next step, so we must have $r(v_i) < r(v_{i+1})$.

This means that the row coordinates never decrease along P, so it is indeed a row-monotone v-u walk. In fact, it is a path because of its minimality. $\qquad\square$

Now if a component of G has vertices in at least $\varepsilon n/18$ rows, then Claims 5.3 and 5.4 together imply that X contains an $\frac{\varepsilon n}{18} \times \frac{\varepsilon n}{18}$ all-1 submatrix, as needed. The next claim shows that if there is no such component, then we can find a large all-1 submatrix in Y, without even forbidding Q.

Claim 5.5. *Suppose no component of G has vertices in $\varepsilon n/18$ different rows. Then Y contains an $\frac{\varepsilon n}{18} \times \frac{\varepsilon n}{18}$ all-1 submatrix.*

Proof. Let $C_1, \ldots, C_k$ be the components of G, and let $r(C_i)$ and $c(C_i)$ be the row and column sets of C_i. Note that all the 0-entries of Y in rows $r(C_i)$ or columns $c(C_i)$ are inside $A[r(C_i) \times c(C_i)]$.

Swapping rows and columns does not affect our statement, so let us reorder the rows and columns of Y so that $r(C_1)$ are the first $|r(C_1)|$ rows, followed by the rows $r(C_2)$, etc., and similarly for columns. This way we get a block-diagonal matrix with blocks $B_i = r(C_i) \times c(C_i)$, where each block has height less than $\varepsilon n/18$ and all the 0-entries are inside the blocks.

Consider the block B_i that touches the $\frac{\varepsilon n}{12}$'th (essentially the middle) column of Y. If no such block exists, then the right half of Y is an $\frac{\varepsilon n}{6} \times \frac{\varepsilon n}{12}$ all-1 submatrix, so we are done. We know that B_i has fewer than $\varepsilon n/18$ rows, so this block cannot contain entries from both the $\frac{\varepsilon n}{18}$'th and the $\frac{2\varepsilon n}{18}$'th rows of Y. If it is disjoint from the $\frac{\varepsilon n}{18}$'th row, then there is an $\frac{\varepsilon n}{18} \times \frac{\varepsilon n}{12}$ all-1 submatrix in the top right corner of Y. Otherwise, we find such a submatrix in the bottom left corner of Y. $\qquad\square$

This completes the proof of Lemma 5.2. $\qquad\square$

Proof of Theorem 2.2. Let A be an $n \times n$ Q-free zero-one matrix. If A contains at least $2n^2/20$ 0-entries, then by Lemma 5.1, it has an $\frac{n}{20} \times \frac{n}{20}$ all-0 submatrix. Otherwise, A contains at least $18n^2/20$ 1-entries, so we can apply Lemma 5.2 to find an $\frac{n}{20} \times \frac{n}{20}$ all-1 submatrix in A. $\qquad\square$

The above proof breaks completely if instead of Q we forbid an arbitrary simple $2 \times k$ matrix, although most of it (including a weakening of Claim 5.3) is salvageable in the special case when we forbid Q_k. Unfortunately, Claim 5.4 is false even in this case, and we do not see any meaningful

way to circumvent it. The best we can do for Q_k-free matrices is to find a homogeneous submatrix of size $\frac{cn}{\log n} \times \frac{cn}{\log n}$ using the methods of Section 6.

6 General $2 \times k$ matrices–Proof of Theorem 2.3

In this section, we prove Theorem 2.3 with the help of partial orders. A *comparability graph* is a graph G whose edges correspond to comparable pairs in some partial order on $V(G)$. The key idea in our proof is to introduce partial orders on the rows of A using the forbidden submatrix. To find the homogeneous submatrices, we need to analyze complete bipartite subgraphs in the comparability graphs and their complement. Our bound on the size of the homogeneous submatrix comes from the following result of Fox and Pach [14].

Theorem 6.1 (Fox, Pach). *Let G be the union of k comparability graphs $G_1, \ldots, G_k$ on the same n vertices. Then either one of the graphs $G_1, \ldots, G_k$ or the complement of G contains a complete bipartite graph with parts of size $n2^{-(1+o(1))(\log \log n)^k}$.*

For simplicity, we write $f_k(n) = n2^{-(1+o(1))(\log \log n)^k}$. We show that if P is $2 \times k$ acyclic, then we can find an all-0 matrix of almost linear size in any P-free zero-one matrix, where the density of 0-entries is positive.

Lemma 6.2. *Let P be an acyclic $2 \times k$ zero-one matrix. For every $\varepsilon > 0$, there is a δ such that every P-free $n \times n$ zero-one matrix with at least εn^2 0-entries contains an $f_k(\delta n) \times \delta n$ all-0 submatrix.*

Proof. Let $\delta = (\frac{\varepsilon n}{16k})^{k+1}$. We will start with some preprocessing on A to find a large submatrix with $k+1$ "nice" all-0 columns, such that every row contains many 0-entries between any two nice columns.

Let us call a $(k+1)$-tuple $(c_1, \ldots, c_{k+1})$ *nice* for a row r if $c_1 < \cdots < c_{k+1}$, $A(r, c_i) = 0$ for every i, and there are at least $\frac{\varepsilon n}{8k}$ 0-entries in $A\big[\{r\} \times [c_i + 1, c_{i+1}]\big]$ for $i = 1, \ldots, k$.

If the r'th row of P contains at least $\frac{\varepsilon n}{2}$ 0-entries, then there are at least $(\frac{\varepsilon n}{8k})^{k+1}$ nice $(k+1)$-tuples for r. Indeed, if the columns of the 0-entries in the r'th row are $j_1 < \cdots < j_\ell$, then every $(k+1)$-tuple $(j_{x_1}, j_{x_2}, \ldots, j_{x_{k+1}})$ is a nice $(k+1)$-tuple for r, whenever $\frac{\varepsilon n i}{4k} - \frac{\varepsilon n}{16k} \leq x_i \leq \frac{\varepsilon n i}{4k} + \frac{\varepsilon n}{16k}$.

The number of rows with at least $\frac{\varepsilon n}{2}$ 0-entries is at least $\frac{\varepsilon n}{2}$. Hence, there are at least $(\frac{\varepsilon n}{8k})^{k+2}$ $(k+1)$-tuples in total (with multiplicities) that are nice for some row. As the number of different $(k+1)$-tuples in $[n]$ is less than n^{k+1}, some $(k+1)$-tuple $(c_1, \ldots, c_{k+1})$ is nice for at least $(\frac{\varepsilon}{8k})^{k+2} n$ rows r. Let V be a set of $(\frac{\varepsilon}{8k})^{k+2} n$ such rows, and let I_i be the interval $[c_i + 1, c_{i+1}]$ for $i = 1, \ldots, k$. Then each row of every matrix $A_i = A[V \times I_i]$ contains at least $\frac{\varepsilon n}{8k}$ 0-entries, and the last column of every A_i is all-0.

For every $i \in [k]$, define the graph G_i on vertex set V as follows. We join a and b in V by an edge if the submatrix of A_i induced by rows $\{a, b\}$ does not contain the i'th column of P. As A is P-free, $\bigcup_{i \in [k]} G_i$ must be the complete graph on V.

Let us make some observations about these graphs. First of all, if the i'th column of P is all-0, then G_i is empty because A_i has an all-0 column. Note also that P can have at most one all-1 column, otherwise it would not be acyclic. Finally (and crucially), if the i'th column of P is not homogeneous, then G_i is a comparability graph. Indeed, suppose that the i'th column of P is $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. For a row $r \in V$, let X_r be the set of columns s such that $A_i(r, s) = 0$. Then for $r, r' \in V$, where $r < r'$, we have that r and r' are joined by an edge in G_i if and only if $X_r \subseteq X_{r'}$. As the relation

10

$\{(r, r') : r < r' \text{ and } X_r \subseteq X_{r'}\}$ is easily seen to be a poset, G_i is indeed a comparability graph. A similar argument works if the i'th column of P is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Let $K \subseteq [k]$ be the set of inhomogeneous columns in P, and let $G = \bigcup_{i \in K} G_i$. By Theorem 6.1, either some G_i or the complement of G contains a complete bipartite graph with parts of size $m = f_k(|V|)$. First suppose that G_i contains $K_{m,m}$ for some $i \in K$. We may assume by symmetry that the i'th column of P is $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Let $v \in V$ be the first row in A_i that appears in this $K_{m,m}$. Then v is adjacent to a set $W \subseteq V$ of m rows below it, and $X_v \subseteq X_w$ for every $w \in W$. Recall that $|X_v| \geq \frac{\varepsilon n}{8k}$ by the construction of A_i, so $A_i[W \times X_i]$ is an $m \times \frac{\varepsilon n}{8k}$ all-0 submatrix of A, as needed.

Now suppose that the complement of G contains $K_{m,m}$. As G_i is empty for all-0 columns of P and $\bigcup_{i \in [k]} G_i$ is the complete graph on V, P must have an all-1 column q, and the $K_{m,m}$ must be a subgraph of G_q. Let $S, T \subseteq V$ be the two vertex classes of this $K_{m,m}$. By the definition of G_q, each column of A_q contains a 1-entry in at most one of $A_q[S \times I_q]$ and $A_q[T \times I_q]$. As A_q has at least $\frac{\varepsilon n}{8k}$ columns, one of $A_q[S \times I_q]$ or $A_q[T \times I_q]$ contains at least $\frac{\varepsilon n}{16k}$ all-0 columns, so A_q has an $m \times \frac{\varepsilon n}{16k}$ all-0 submatrix, finishing the proof. $\qquad\square$

Proof of Theorem 2.3. Let A be an $n \times n$ P-free matrix. As P is simple, both P and P^c are acyclic. So, if A has at least $n^2/2$ 0-entries, we can apply Lemma 6.2 to A with P and $\varepsilon = 1/2$. Otherwise, we can apply the lemma to A^c with P^c and $\varepsilon = 1/2$. Either way, we find an $n^{1-o(1)} \times \Omega(n)$ homogeneous submatrix in A. $\qquad\square$

Note that any improvement in Theorem 6.1 would also improve our theorem. However, this alone will not be sufficient to find a linear-size homogeneous submatrix. Indeed, as was shown recently by Korándi and Tomon [24], the size of the bipartite graph in Theorem 6.1 cannot be replaced by anything larger than $\Omega(n/(\log n)^k)$.

On the other hand, one can find slightly larger all-0 submatrices in Lemma 6.2 by reducing the number of partial orders we use. For example, we may assume that K in the proof has size at most $k-1$, as otherwise there are no homogeneous columns in P and we can apply Theorem 2.1. This immediately guarantees an $f_{k-1}(\delta n) \times \delta n$ homogeneous submatrix.

It is also enough to use just one matrix A_i (and comparability graph G_i) for consecutive columns of P if they are the same. For example, if ℓ consecutive columns equal $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, then one can take G_i to be the comparability graph where two rows $r < r'$ are joined by an edge if $|X_r \setminus X_{r'}| \leq (\ell-1)(r-r')$, and use it to embed all ℓ columns in A_i. With this argument, one can find a $\Omega(\frac{n}{\log n}) \times \Omega(n)$ homogeneous submatrix in any $n \times n$ Q_k-free zero-one matrix.

7 Matrices without two ones in a column–Proof of Theorem 2.6

In this section, we prove Theorem 2.6. The main part of our proof is to prove a weaker variant of Conjecture 2.4 for zero-one matrices P with no more than one 1-entry per column. Namely, we show that for some $\varepsilon > 0$, every P-free $n \times n$ matrix A with at least $(1 - \varepsilon)n^2$ 0-entries contains an $\varepsilon n \times \varepsilon n$ all-0 submatrix. This will be enough to obtain Theorem 2.6 when A is very dense or very sparse in terms of 0-entries. For the range in between, we will use the following result of Alon, Fischer, and Newman [1].

Lemma 7.1 (Alon, Fischer, Newman). *Let P be a zero-one matrix. For every $\varepsilon > 0$ there is a $\delta > 0$ such that every P-free $n \times n$ zero-one matrix A has a $\delta n \times \delta n$ submatrix B that has either at most $\varepsilon(\delta n)^2$ or at least $(1 - \varepsilon)(\delta n)^2$ 0-entries.*

Lemma 7.1 is stated in [1, Lemma 1.6] in a much stronger form in a "removal lemma"-type setting, with strong quantitative bounds on δ. However, this weak corollary already serves our purposes. Also, let us remark that in the graph world, this lemma corresponds to the well known result of Rödl [33] that for any graph H, induced H-free graphs cannot have a uniform edge distribution.

Lemma 7.2. *Let P be a zero-one matrix such that no column of P contains more than one 1-entry. Then there is an $\varepsilon = \varepsilon(P) > 0$ such that every P-free $n \times n$ zero-one matrix A with at least $(1-\varepsilon)n^2$ 0-entries has an $\varepsilon n \times \varepsilon n$ all-0 submatrix.*

Proof. Suppose P has $k-1$ rows and ℓ columns. Let I be the $k \times k$ identity matrix, and let R be the $k \times (k\ell)$ matrix that is the concatenation of ℓ copies of I, i.e., $R(i, i+jk) = 1$ for every $i = 1, \ldots, k$ and $j = 0, \ldots, \ell-1$, and all other entries of R are 0. It is easy to see that R contains every $(k-1) \times \ell$ matrix with at most one 1-entry per column as a submatrix.[2] In particular, every P-free matrix is also R-free, so it is enough to prove our theorem for R instead of P.

Let $s = 2(\ell-1)(2k)^k$, $m = \frac{n}{s}$ and $\varepsilon = \frac{1}{8s^2k^2}$. We will show that if A contains at least $(1-\varepsilon)n^2$ 0-entries but does not have an $\varepsilon n \times \varepsilon n$ all-0 submatrix, then A contains R as a submatrix.

Let us split the first m rows of A into $m \times m$ submatrices $A_i = A\big[[m] \times [(i-1)m+1, im]\big]$ for $i \in [s]$. Let T_i be the family of k-element sets $S \subseteq [m]$ such that A_i contains a copy of I in the rows indexed by S.

Claim 7.3. $|T_i| \geq \frac{1}{2}\left(\frac{m}{2k}\right)^k$ *for every $i \in [s]$.*

Proof. Let us consider the matrices

$$A_{i,j} = A_i\left[\left[\frac{(j-1)m}{k}+1, \frac{jm}{k}\right] \times \left[\frac{(j-1)m}{k}+1, \frac{jm}{k}\right]\right]$$

for every $j \in k$. Then $A_{i,1}, \ldots, A_{i,k}$ are $\frac{m}{k} \times \frac{m}{k}$ submatrices along the diagonal of A_i.

As $\frac{m}{2k} \geq \sqrt{\varepsilon}n$, we know that A_i does not have any $\frac{m}{2k} \times \frac{m}{2k}$ all-0 submatrix. This easily implies that in each $A_{i,j}$, there are at least $\frac{m}{2k}$ 1-entries such that no two share a row or a column. Let $S_{i,j}$ be the set of coordinates of these $\frac{m}{2k}$ 1-entries.

Let us pick an element $(x_j, y_j) \in S_{i,j}$ for every $j = 1, \ldots, k$ (so one 1-entry from each $A_{i,j}$), and consider the $k \times k$ submatrix $B = A_i[\{x_1, \ldots, x_k\} \times \{y_1, \ldots, y_k\}]$. There are $\left(\frac{m}{2k}\right)^k$ such submatrices B. Also, B has 1-entries in the diagonal, so $B = I$, unless there is another 1-entry in B. However, each such 1-entry of A_i can appear in at most $\left(\frac{m}{2k}\right)^{k-2}$ matrices B, because it fixes the choice of (x_j, y_j) for two j's: if $A_i(x, y) = 1$, then the 1-entry at (x, y) can only appear in matrices B for which $x_a = x$ for $a = \lceil x/k \rceil$ and $y_b = y$ for $b = \lceil y/k \rceil$. As there are at most εn^2 1-entries in A, we are left with at least

$$\left(\frac{m}{2k}\right)^k - \varepsilon n^2\left(\frac{m}{2k}\right)^{k-2} = \left(\frac{m}{2k}\right)^k - \frac{m^2}{8k^2}\left(\frac{m}{2k}\right)^{k-2} = \frac{1}{2}\left(\frac{m}{2k}\right)^k$$

choices where $B = I$. $\qquad \square$

Suppose that A does not contain R as a submatrix. Then every k-element set $S \subseteq [m]$ can appear in at most $\ell - 1$ of the sets $T_1, \ldots, T_s$. Indeed, if $S \in T_{i_1} \cap T_{i_2} \cap \cdots \cap T_{i_\ell}$, then A contains R as a submatrix in the rows induced by S.

Together with Claim 7.3, this gives

$$\frac{s}{2}\left(\frac{m}{2k}\right)^k \leq \sum_{i=1}^s |T_i| \leq (\ell-1)\binom{m}{k} < (\ell-1)m^k.$$

---

[2] In fact, they are already contained in the first $k-1$ rows of R. We use R for the sake of a simpler presentation.

This contradicts our choice of s. $\qquad\qquad\square$

Proof of Theorem 2.6. By Lemma 7.2, there is an $\varepsilon > 0$ such that any P-free $n \times n$ zero-one matrix with at least $(1-\varepsilon)n^2$ 0-entries contains an $\varepsilon n \times \varepsilon n$ all-0 submatrix. We can apply Lemma 7.1 with this ε to get some $\delta > 0$ such that any P-free $n \times n$ zero-one matrix has a $\delta n \times \delta n$ submatrix B with at least $(1-\varepsilon)(\delta n)^2$ entries that are all 0 or all 1.

Let A be an $n \times n$ zero-one matrix that is both P-free and P^c-free, and let B be the $\delta n \times \delta n$ submatrix with at least $(1-\varepsilon)(\delta n)^2$ equal entries. If these entries are all 0, then B contains an $\varepsilon \delta n \times \varepsilon \delta n$ all-0 submatrix because it is P-free. Otherwise, B^c is a P-free matrix with at least $(1-\varepsilon)(\delta n)^2$ 0-entries, so B contains an $\varepsilon \delta n \times \varepsilon \delta n$ all-1 submatrix. $\qquad\square$

8 Unordered matrices–Proof of Theorem 1.6

In this section, we prove Theorem 1.6. Again, we show that if an unordered P-free matrix has a positive density of 0-entries, then it contains a linear-size all-0 submatrix.

Lemma 8.1. *Let P be a simple $2 \times k$ zero-one matrix. Then every unordered P-free $n \times n$ zero-one matrix with at least εn^2 0-entries contains an $\frac{\varepsilon n}{6k} \times \frac{\varepsilon n}{6k}$ all-0 submatrix.*

Proof. Let R be the $2 \times (2k+2)$ matrix whose first k columns are $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, the next k columns are $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, the $(2k+1)$'st column is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and the last column is $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Then R contains an ordering of the columns of P, so it is enough to prove our result for R instead of P.

Let A' be the matrix obtained from A by deleting the rows with fewer than $\varepsilon n/2$ 0-entries. At most $\varepsilon n^2/2$ 0's are deleted, so A' contains at least $\varepsilon n^2/2$ 0-entries. Hence, one can find a column in A' with $t = \lceil \varepsilon n/2 \rceil$ 0 entries. Let B be the $t \times n$ submatrix of A induced by the rows of these 0-entries.

By permuting rows and columns if necessary, we may assume that these 0-entries form an all-0 last column in B, and that the rows of B are in increasing order according to the number of 0-entries in them.

For $i \in [t]$, let H_i denote the set of indices $j \in [n]$ such that $B(i,j) = 0$. Define the directed graph G on vertex set $[t]$ by adding (i,j) as an edge if $i < j$ and $|H_i \setminus H_j| \leq k - 1$. Then G is an acyclic directed graph.

Note that if (i,j) is not an edge of G for some $i < j$, then we must have $H_i \cup H_j = [n]$. Indeed, if $r \in [n] \setminus (H_i \cup H_j)$, then $B[\{i,j\} \times \{r\}] = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. We also have $|H_i \setminus H_j| \geq k$, which further implies $|H_j \setminus H_i| \geq k$ because $|H_i| \leq |H_j|$. Therefore, if X is a k-element subset of $H_i \setminus H_j$, and Y is a k-element subset of $H_j \setminus H_i$, then $B[\{i,j\} \times (X \cup Y \cup \{r,n\})]$ is a reordering of R, contradicting our assumption.

For a set $Z \subseteq [n]$, we denote its complement by $\overline{Z} = [n] \setminus Z$. Let M be the set of minimal vertices in G, that is, the set of vertices v such that no edge points towards v. Then the sets $\overline{H}_v$ are pairwise disjoint for $v \in M$. Every element $w \in [t]$ can be reached from a minimal vertex via a directed path. Let us assign each w to the one such vertex in M with the smallest label in $[t]$.

Now we will show that there is a subset $N \subseteq M$ such that $|\bigcup_{v \in N} \overline{H}_v| \leq n - t$ and at least $t/3$ of the elements in $[t]$ are assigned to vertices in N. Note that by the construction of B, we have $|H_i| \geq t$ and hence $|\overline{H}_i| \leq n - t$ for every $i \in [t]$. If M contains a vertex v that is assigned to more than $t/3$ elements of $[t]$, then we are done, as we can take $N = \{v\}$. So we may assume that there

13

is no such vertex. Starting with $N_0 = \emptyset$, add vertices of M one by one to N_0 until the number of elements assigned to the vertices in N_0 is at least $t/3$. At this point, the number of elements assigned to N_0 is between $t/3$ and $2t/3$. If $|\bigcup_{v \in N_0} \overline{H}_v| \leq n - t$, then set $N = N_0$, otherwise, set $N = M \setminus N_0$. As $t = \lceil \varepsilon n/2 \rceil \leq \lceil n/2 \rceil$, we must have $|\bigcup_{v \in N} \overline{H}_v| \leq n - |\bigcup_{v \in N_0} \overline{H}_v| \leq t - 1 \leq n - t$. The number of elements assigned to an element of N is at least $t/3$ in both cases.

Now let $x_1 < \cdots < x_s$ be the elements of $[t]$ assigned to N, so $s \geq t/3$. Also, for $X = \bigcap_{v \in N} H_v$, we have $|X| \geq t$. We show by induction on ℓ that $|X \cap \bigcap_{i=1}^{\ell} H_{x_i}| \geq |X| - (\ell - 1)(k - 1)$. If $\ell = 1$, then x_1 is a minimal element, so $X \subseteq H_{x_1}$, and we are done. Now suppose that $\ell > 1$. If $x_\ell \in N$, then $X \subseteq H_{x_\ell}$, so

$$\left| X \cap \bigcap_{i=1}^{\ell} H_{x_i} \right| = \left| X \cap \bigcap_{i=1}^{\ell-1} H_{x_i} \right| \geq |X| - (\ell - 2)(k - 1) \geq |X| - (\ell - 1)(k - 1),$$

and we are done. If $x_\ell \notin N$, then G must contain an edge $(x_{\ell'}, x_\ell)$ for some $1 \leq \ell' < \ell$. Indeed, if x_ℓ is assigned to $v \in N$, then all other vertices on a v-x_ℓ directed path are assigned to v, as well. Now we can use $|H_{x_{\ell'}} \setminus H_{x_\ell}| \leq k - 1$, and hence $|(X \cap \bigcap_{i=1}^{\ell-1} H_{x_i}) \setminus H_{x_\ell}| \leq k - 1$, to get

$$\left| X \cap \bigcap_{i=1}^{\ell} H_{x_i} \right| \geq \left| X \cap \bigcap_{i=1}^{\ell-1} H_{x_i} \right| - (k - 1) \geq |X| - (\ell - 1)(k - 1).$$

Fix $\ell = \min\{\frac{t}{2(k-1)}, \frac{t}{3}\}$ (for $k = 1$, take $\ell = \frac{t}{3}$). Then $|\bigcap_{i=1}^{\ell} H_{x_i}| \geq |X| - (\ell - 1)(k - 1) \geq t/2$, so the submatrix of B induced by the rows $\{x_1, \ldots, x_\ell\}$ and columns $\bigcap_{i=1}^{\ell} H_{x_i}$ is an all-0 matrix with at least $\min\{\frac{t}{2(k-1)}, \frac{t}{3}\} \geq \frac{\varepsilon n}{6k}$ rows, and at least $\frac{t}{2} \geq \frac{\varepsilon n}{4}$ columns. This finishes the proof. $\qquad\square$

Proof of Theorem 1.6. If A has at least $\frac{n^2}{2}$ 0-entries, we can find a $\frac{n}{12k} \times \frac{n}{12k}$ all-0 submatrix in A by the previous lemma. Otherwise, we can apply Lemma 8.1 to A^c to show that A contains a $\frac{n}{12k} \times \frac{n}{12k}$ all-1 submatrix. $\qquad\square$

Lemma 8.1 shows that there is a genuine difference between the ordered and unordered case of our problem. Indeed, this result shows that εn^2 0-entries in an *unordered* P-free matrix guarantee an $\Omega(\varepsilon n) \times \Omega(\varepsilon n)$ all-0 submatrix. However, as we discussed at the end of Section 4, this is not true for every $2 \times k$ matrix P in the ordered setting: there are P-free matrices with εn^2 0-entries that do not have any all-0 submatrix of size $\Omega(\frac{\varepsilon}{\alpha(1/\varepsilon)} n)$.

By a result of Füredi [17], there is an $n \times n$ matrix A with $\Theta(n \log n)$ 0-entries that does not contain $\begin{pmatrix} 0 & * & 0 \\ 0 & 0 & * \end{pmatrix}$ (where $*$ can be either 1 or 0). With the same methods as before, we can use this to construct $n \times n$ matrices A with εn^2 0-entries that do not contain $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and have no all-0 submatrices of size $\Omega(\frac{\varepsilon}{\log 1/\varepsilon} n)$.

9   Applications

Several matrix classes can be described by a finite set of forbidden submatrices (see, e.g., [22]), and our results show that in many cases they contain large homogeneous submatrices. We give three specific applications.

9.1 Chordal bipartite graphs and totally balanced matrices

A zero-one matrix is *totally balanced* if it does not contain any submatrix, whose columns are different and which has exactly two 1-entries in each of its rows and columns. In other words, none of its submatrices is the incidence matrix of a cycle of length at least 3.

Totally balanced matrices (first studied by Lovász [26] in connection with a hypergraph coloring problem) are well-examined objects in combinatorial optimization. Their importance comes from the fact that integer programs with totally balanced coefficient matrices can be easily solved. Indeed, the optimization problem can be solved greedily if the coefficient matrix does not contain $\Gamma = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ as a submatrix, and as was shown in [5, 20, 28], a matrix is totally balanced if and only if its rows and columns can be rearranged to get a Γ-free matrix. (For more on optimization properties of balanced matrices, see the book of Berge [7].) As rearranging rows and columns does not affect homogeneous submatrices, Theorem 2.2 shows that totally balanced matrices have large homogeneous submatrices.

Corollary 9.1. *Every totally balanced $n \times n$ matrix contains an $cn \times cn$ homogeneous submatrix with some $c \geq 1/20$.*

A *chordal bipartite graph* is a bipartite graph with no induced cycle of length greater than 4. This class of graphs was introduced by Golumbic and Goss [19] as a bipartite analog to chordal graphs, with similar perfect elimination properties. Clearly, a bipartite graph is chordal if and only if its adjacency matrix is totally balanced. This immediately implies the following.

Corollary 9.2. *Every chordal bipartite graph $G = (A \cup B, E)$ with parts of size n contains sets $A' \subseteq A$ and $B' \subseteq B$ of size cn, for some constant $c > 0$, such that $G[A', B']$ is either empty or complete.*

9.2 The Erdős-Hajnal conjecture and intersection graphs

A family $\mathcal{G}$ of graphs is said to have the *Erdős-Hajnal property*, if there is a constant c such that each member $G \in \mathcal{G}$ contains a clique or an independent set on at least $|V(G)|^c$ vertices. The family $\mathcal{G}$ has the *strong Erdős-Hajnal property*, if there is a constant c' such that every $G \in \mathcal{G}$ satisfies that either G or its complement contains a complete bipartite graph with parts of size $c'|V(G)|$. By a result of Alon, Pach, Pinchasi, Radoičić and Sharir [3], the strong Erdős-Hajnal property implies the Erdős-Hajnal property in hereditary families. The famous Erdős-Hajnal conjecture [12, 13] asserts the following.

Conjecture 9.3 (Erdős, Hajnal). *For every graph H, the family of graphs not containing an induced copy of H has the Erdős-Hajnal property.*

This conjecture has attracted significant attention in the past decades, but is still wide open. For history and relevant results, we refer the reader to the survey of Chudnovsky [10].

The *intersection graph* of a family of sets $\mathcal{F}$ is the graph with vertex set $\mathcal{F}$, where two vertices are joined by an edge if their intersection is nonempty. A *curve* in the plane is the image of an injective continuous function $f : [0, 1] \to \mathbb{R}^2$. In this paper, we assume that curves in our collections only meet at proper crossings, that is, if two curves α and β share a point in common, then α passes to the other side of β at this point. A *string graph* is a graph that is isomorphic to the intersection graph of a family of curves.

In a very recent paper, Tomon [35] showed that the family of string graphs has the Erdős-Hajnal property. However, this family does not satisfy the strong Erdős-Hajnal property [32], although Fox

and Pach [15] proved that one can always find a complete bipartite graph of almost linear size in every string graph or its complement.

Theorem 9.4 (Fox, Pach). *Let G be a string graph on n vertices. Then either G contains $K_{m,m}$ with $m = \Omega(\frac{n}{\log n})$, or the complement of G contains $K_{m',m'}$ with $m' = \Omega(n)$.*

A collection of curves is *k-intersecting*, if any two curves in the collection intersect in at most k points. Fox, Pach and Tóth [16] showed that the family of intersection graphs of k-intersecting curves does have the strong Erdős-Hajnal property.

Theorem 9.5 (Fox, Pach, Tóth). *For every positive integer k, there is a constant $c_k > 0$ such that the following holds. Let G be the intersection graph of a k-intersecting family of n curves. Then either G or its complement contains a complete bipartite graph of size $c_k n$.*

Here, we are interested in a bipartite version of this problem. That is, given two families of n curves, $\mathcal{A}$ and $\mathcal{B}$, we would like to find large subfamilies, $\mathcal{A}_0 \subseteq \mathcal{A}$ and $\mathcal{B}_0 \subseteq \mathcal{B}$, such that $|\mathcal{A}_0| = |\mathcal{B}_0|$, and either every curve in $\mathcal{A}_0$ intersects every curve in $\mathcal{B}_0$, or every curve in $\mathcal{A}_0$ is disjoint from every curve in $\mathcal{B}_0$.

In general, we cannot hope for any bound on $|\mathcal{A}_0| = |\mathcal{B}_0|$ beating the Ramsey bound $\Theta(\log n)$. Indeed, the complement of every comparability graph is a string graph [27, 32], therefore the complement of any bipartite graph is a string graph. Nevertheless, the question remains meaningful if we restrict ourselves to k-intersecting collections of curves.

In fact, we believe that the condition that $\mathcal{A} \cup \mathcal{B}$ is k-intersecting can be weakened to only requiring that $\mathcal{A}$ and $\mathcal{B}$ themselves are k-intersecting.

Conjecture 9.6. *For every k there is a constant $c_k > 0$ such that the following holds. Let $\mathcal{A}$ and $\mathcal{B}$ be two families of n curves each such that $\mathcal{A}$ and $\mathcal{B}$ are k-intersecting. Then there are subfamilies $\mathcal{A}_0 \subseteq \mathcal{A}$ and $\mathcal{B}_0 \subseteq \mathcal{B}$ such that $|\mathcal{A}_0| = |\mathcal{B}_0| \geq c_k n$, and either every $\alpha \in \mathcal{A}_0$ intersects every $\beta \in \mathcal{B}_0$, or every $\alpha \in \mathcal{A}_0$ is disjoint from every $\beta \in \mathcal{B}_0$.*

In some sense, this is the weakest condition one can impose on $\mathcal{A}$ and $\mathcal{B}$ to force any meaningful properties. Indeed, the complement of any bipartite graph can be realized as the intersection graph of a collection of curves $\mathcal{A} \cup \mathcal{B}$, where $\mathcal{A}$ is 1-intersecting, and any two curves $A \in \mathcal{A}$ and $B \in \mathcal{B}$ intersect in at most 2 points (but $\mathcal{B}$ is not k-intersecting for any bounded k), see [31].

A natural special case of the conjecture is when the curves are 0-1 curves. Here, a *0-1 curve* is the drawing of a continuous function $f : [0,1] \to \mathbb{R}$ in $\mathbb{R}^2$. As a first step towards Conjecture 9.6, we prove the following statement.

Theorem 9.7. *Let $\mathcal{A}$ and $\mathcal{B}$ be two families of n 0-1 curves each. If $\mathcal{A}$ is k-intersecting, and $\mathcal{B}$ is 1-intersecting, then there are subfamilies $\mathcal{A}_0 \subseteq \mathcal{A}$ and $\mathcal{B}_0 \subseteq \mathcal{B}$ such that $|\mathcal{A}_0| = |\mathcal{B}_0| \geq \Omega(n/k)$, and either every $\alpha \in \mathcal{A}_0$ intersects every $\beta \in \mathcal{B}_0$, or every $\alpha \in \mathcal{A}_0$ is disjoint from every $\beta \in \mathcal{B}_0$.*

Proof. By slightly perturbing our curves, we can assume that no 3 curves in $\mathcal{A} \cup \mathcal{B}$ go through the same point, and no two of them intersect the lines $x = 0$ and $x = 1$ in the same point. For two curves $\gamma, \gamma' \in \mathcal{A} \cup \mathcal{B}$, let $\gamma \prec \gamma'$ if γ intersects the vertical line $x = 0$ below γ'.

First, we claim that there are subfamilies $\mathcal{A}' \subseteq \mathcal{A}$ and $\mathcal{B}' \subseteq \mathcal{B}$ such that $|\mathcal{A}'| = |\mathcal{B}'| = \lceil n/2 \rceil$, and either $\alpha \prec \beta$ for every $(\alpha, \beta) \in \mathcal{A}' \times \mathcal{B}'$, or $\beta \prec \alpha$ for every $(\alpha, \beta) \in \mathcal{A}' \times \mathcal{B}'$. Indeed, in the total ordering defined by $\prec$, pick the smallest element $\gamma \in \mathcal{A} \cup \mathcal{B}$ such that either $\lceil n/2 \rceil$ elements of $\mathcal{A}$ are $\preceq \gamma$, or $\lceil n/2 \rceil$ elements of $\mathcal{B}$ are $\preceq \gamma$. In the first case, set $\mathcal{A}' = \{\alpha \in \mathcal{A} : \alpha \preceq \gamma\}$ and let $\mathcal{B}'$ be an $\lceil n/2 \rceil$ element subset of $\{\beta \in \mathcal{B} : \gamma \prec \beta\}$. In the second case, let $\mathcal{A}'$ be an $\lceil n/2 \rceil$ element subset of $\{\alpha \in \mathcal{A} : \gamma \prec \alpha\}$ and $\mathcal{B}' = \{\beta \in \mathcal{B} : \beta \preceq \gamma\}$.
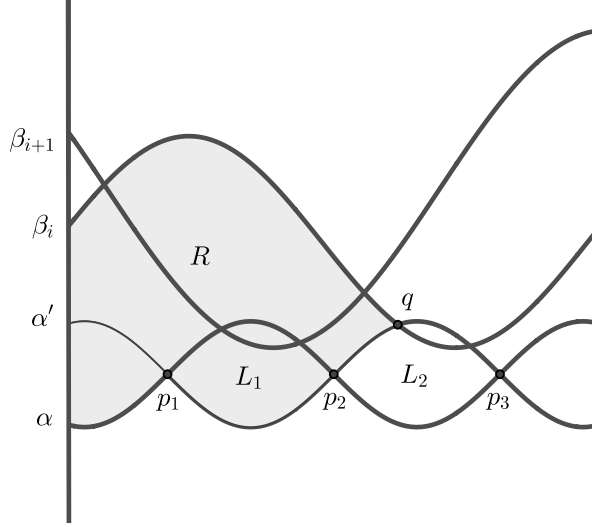
Figure 1: An illustration of the proof of Claim 9.8

Without loss of generality, suppose that $\alpha \prec \beta$ for every $(\alpha, \beta) \in \mathcal{A}' \times \mathcal{B}'$. Define the $\lceil n/2 \rceil \times \lceil n/2 \rceil$ matrix A by setting $A(i,j) = 1$ if the i'th smallest element of $\mathcal{A}'$ intersects the j'th smallest element of $\mathcal{B}'$ with respect to the ordering $\prec$, and $A(i,j) = 0$ otherwise.

Claim 9.8. *Let P_ℓ be the $2 \times \ell$ matrix defined by $P_\ell(i,j) = 1$, if $i + j$ is even, and $P_\ell(i,j) = 0$ if $i + j$ is odd. Then A is P_{k+2}-free.*

Proof. Let us start with introducing some notation. Each 0-1 curve cuts the strip $[0,1] \times \mathbb{R}$ into two parts, an upper and lower part. We say that a point set is *above* the curve if it is a subset of the upper part, and it is *below*, if it is a subset of the lower part. Also, if γ is a 0-1 curve and $q \in \gamma$, let $\gamma(q)$ denote the subcurve of γ starting on the vertical line $x = 0$, and ending at q. For $q, q' \in \gamma$, we define $\gamma(q, q') = \gamma(q') \setminus \gamma(q)$.

Suppose that $\alpha \prec \alpha'$ in $\mathcal{A}$, and $\beta_1 \prec \cdots \prec \beta_{k+2}$ in $\mathcal{B}$ induce P_{k+2}. Let $p_1, \ldots, p_t$ be the intersection points of the curves α and α', ordered by their x-coordinates. As $\mathcal{A}$ is k-intersecting, we have $t \leq k$. These t intersection points cut both α and α' into $k + 1$ subcurves, let us denote them by $\alpha_0, \ldots, \alpha_t$ and $\alpha'_0, \ldots, \alpha'_t$ from left to right. Note that $\alpha < \alpha'$ implies that if i is even, then α_i is below α', and α'_i is above α, while if i is odd, then α_i is above α' and α'_i is below α. For $i = 0, \ldots, t$, let L_i denote the region in $[0,1] \times \mathbb{R}$ bounded by α_i and α'_i, and call these regions L_i *lenses*. If i is even, say that α'_i is the *top boundary* of L_i and α_i is the *bottom boundary*, and if i is odd, then α_i is the top boundary of L_i, and α'_i is the bottom boundary. Note that if β_j intersects the lens L_i, then β_j intersects only the top boundary of L_i, as $\alpha, \alpha' \prec \beta_i$ and each of the curves β_i intersect exactly one of α and α'. Therefore, if L_i and β_j intersect, i and j must have the same parity.

For $i \in [k+2]$, let $\ell(i)$ denote the smallest index for which β_i intersects the lens $L_{\ell(i)}$. We show that $\ell(1) < \ell(2) < \cdots < \ell(k + 2)$, which contradicts $0 \leq \ell(i) \leq k$. Suppose that $\ell(i + 1) \leq \ell(i)$ for some $i \in [k + 1]$. As i and $i + 1$ have different parities, we have $\ell(i + 1) < \ell(i)$ and β_{i+1} cannot intersect $L_{\ell(i)}$. Let γ denote the union of the top boundaries of all the lenses, and let γ' denote the union of the bottom boundaries of the lenses, then γ and γ' are 0-1 curves. Let q be the first intersection point of γ and β_i, and let

$$\delta = \gamma'(p_{\ell(i)}) \cup \gamma(p_{\ell(i)}, q).$$

17

In other words, we obtain the curve δ by following the bottom boundaries of the lenses until we reach the lens $L_{\ell(i)}$, where we follow the top boundary until we reach β_i. Let R be the region bounded by $\beta_i(q)$ and δ, see Figure 1. The curve β_{i+1} starts outside R, but R contains the lens $L_{\ell(i+1)}$, so β_{i+1} must enter R. However, β_{i+1} does not intersect intersect γ', nor does it touch $L_{\ell(i)}$. Thus, β_{i+1} cannot intersect δ. Hence, β_{i+1} must enter R through $\beta_i(q)$. Since β_{i+1} also leaves R, it must also exit through $\beta_i(q)$. Therefore, β_i and β_{i+1} intersect twice, contradiction. $\qquad\square$

The $2 \times k$ matrix P_{k+2} does not contain a homogeneous column, so we can apply Theorem 2.1 to conclude that A contains a homogeneous submatrix of size at least $\Omega(n/k)$. This corresponds to two collections $\mathcal{A}_0 \subseteq \mathcal{A}$ and $\mathcal{B}_0 \subseteq \mathcal{B}$ with the desired properties.

$\qquad\square$

9.3 Pseudohalfplanes

A *bi-infinite x-monotone curve* is the graph of a continuous function $f : \mathbb{R} \to \mathbb{R}$. A collection $\mathcal{L}$ of bi-infinite x-monotone curves is a *pseudoline-arrangement* if any two elements of $\mathcal{L}$ intersect in exactly one point. If $\mathcal{L}$ is a pseudoline-arrangement, then $\mathcal{H}$ is a *pseudohalfplane-arrangement* if every element $H \in \mathcal{H}$ is either the set of points below an element of $\mathcal{L}$, or the set of points above an element of $\mathcal{L}$.

Let P be a set of points in the plane and let $\mathcal{H}$ be a pseudohalfplane-arrangement. Consider the matrix M whose rows are labeled with elements of P, columns are labeled with the elements of $\mathcal{H}$, and

$$M(p, H) = \begin{cases} 1 & \text{if } p \in H \\ 0 & \text{if } p \notin H \end{cases}.$$

It is proved in [21, Theorem 2.19, Proposition A.1] (see also [9]) that M can be partitioned into two submatrices M_1 and M_2 such that the following holds: the rows and columns of M_1 and M_2 can be ordered such that M_1 and M_2 does not contain $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ as a submatrix. But then Theorem 1.1 immediately implies that some linear set of pseudohalfplanes contains or avoids a positive proportion of the points.

Corollary 9.9. *Let P be a set of n points in the plane and let $\mathcal{H}$ be a pseudohalfplane-arrangement with n elements. Then there are subsets $P_0 \subset P$ and $\mathcal{H}_0 \subset \mathcal{H}$ of size $|P_0| = |\mathcal{H}_0| \geq cn$ for a suitable constant $c > 0$, such that either for every $p \in P_0$ and $H \in \mathcal{H}_0$ we have $p \in H$, or for every $p \in P_0$ and $H \in \mathcal{H}_0$ we have $p \notin H$.*

10 Concluding remarks

Our work establishes various bounds on the size of the largest homogeneous submatrix that can be found in a matrix, when a fixed submatrix P is forbidden. A summary of our results for fixed small P can be found in Figure 2. A number of questions remain unsolved, and it would be very interesting to obtain good bounds for simple or acyclic matrices. Perhaps the first open question is to decide if $Q_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ satisfies Conjecture 2.5, i.e., if forbidding the submatrix Q_2 in an $n \times n$ zero-one matrix guarantees the existence of a $cn \times cn$ homogeneous submatrix.

These questions are also closely related to recent results on the Erdős-Hajnal theory of trees: Extending previous work in [8, 25], Chudnovsky, Scott, Seymour and Spirkl [11] proved the following variant of the Erdős-Hajnal conjecture.

18

P	Conjecture 2.4	Best bound	Reference
$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	✓	$\geq cn \times cn$	Lemma 3.4
$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	✓	$\geq cn \times cn$	Lemma 5.1
$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$	✓	$\geq cn \times cn$	Lemma 4.2
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	✓	$\geq cn \times cn$	Lemma 4.2
$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$	✓	$\geq cn \times cn$	Lemma 5.2
$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	✗	$\leq n^{1-\varepsilon} \times n^{1-\varepsilon}$	Proposition 3.1
$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	✓	$\geq cn \times cn$	Lemma 5.1
$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$	✓	$\geq cn \times cn$	Lemma 4.2
$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$	?	$\geq \frac{cn}{\log n} \times cn$	Lemma 6.2
$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$	?	$\geq n^{1-o(1)} \times cn$	Lemma 6.2
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	?	$\geq cn \times cn$ if $\geq (1-\varepsilon)n^2$ 0-entries	Lemma 7.2

Figure 2: Largest size of an all-0 submatrix in an $n \times n$ P-free matrix with at least εn^2 0-entries.

Theorem 10.1 (Chudnovsky et al.). *Let T be a tree. Then the family of all graphs not containing an induced copy of T and T^c has the strong Erdős-Hajnal property.*

Our problems can be thought of as an ordered bipartite version of the strong Erdős-Hajnal problem. For example, it is not hard to see that Theorem 10.1 would follow from Conjecture 2.5.

Indeed, we can think of our $n \times n$ zero-one matrix A as the biadjacency matrix of a bipartite graph $G(A \cup B, E)$ with parts of size n. Submatrices then correspond to induced subgraphs, and a homogeneous submatrix means a subgraph $G' = (A' \cup B', E')$ that is complete or empty between $A' \subseteq A$ and $B' \subseteq B$. An important difference, though, is that a forbidden submatrix only forbids *one ordering* of the corresponding bipartite graph (where the vertices in the two parts are ordered according to the rows and columns of the matrix). This is a much weaker condition and adds considerable difficulty to our problem.

Approximate versions of our Conjectures 2.4 and 2.5, finding $n^{1-o(1)} \times n^{1-o(1)}$ homogeneous submatrices, were very recently proved by Scott, Seymour and Spirkl [34].

Extremal questions about zero-one matrices have been extensively studied over the past decades, and it is worth mentioning a few that are loosely related to our problem.

A zero-one matrix A contains a *pattern P*, where P is another zero-one matrix, if P can be

obtained from a submatrix of A by changing some 1-entries to 0-entries. When A is a biadjacency matrix, this corresponds to the subgraph relation (as opposed to submatrices corresponding to *induced* subgraphs). The Turán number $\text{ex}(n, P)$ is defined as the maximum number of 1-entries in an $n \times n$ zero-one matrix that does not contain the pattern P. A central problem in this area is a conjecture of Pach and Tardos [30] that $\text{ex}(n, P) = O(n \operatorname{polylog} n)$ whenever P is acyclic. Although this is known for many such matrices [18, 29, 30, 23], the general conjecture remains open.

Another related question asks for $\text{forb}(n, P)$, the maximum number of distinct columns in an unordered P-free zero-one matrix A with n rows. When we think of A as the incidence matrix of a hypergraph, finding $\text{forb}(n, P)$ is connected to certain hypergraph coloring problems (see, e.g., [26]), as well as other structural results. For example, in the special case when P is the $k \times 2^k$ zero-one matrix with all different columns, the Sauer-Shelah lemma gives $\text{forb}(n, P) = \binom{n}{k-1} + \binom{n}{k-2} + \cdots + \binom{n}{0}$. An open conjecture of Anstee and Sali [6] asserts that $\text{forb}(n, P) = \Theta(n^{f(P)})$ for an implicitly defined integer function f. For further partial results on this topic, we refer the reader to the survey [4].

Acknowledgments

References

[1] N. Alon, E. Fischer, I. Newman, *Efficient testing of bipartite graphs for forbidden induced subgraphs,* SIAM Journal on Computing 37 (2007): 959–976.

[2] M. Axenovich, C. Tompkins, L. Weber, *Large homogeneous subgraphs in bipartite graphs with forbidden induced subgraphs,* arXiv:1903.09725 preprint.

[3] N. Alon, J. Pach, R. Pinchasi, R. Radoičić, M. Sharir, *Crossing patterns of semi-algebraic sets,* Journal of Combinatorial Theory Series A 111 (2) (2005): 310–326.

[4] R.P. Anstee, *A survey of forbidden configuration results,* Electronic Journal of Combinatorics DS20 (2013): pp53.

[5] R.P. Anstee, M. Farber, *Characterizations of totally balanced matrices,* Journal of Algorithms 5 (1984): 215–230.

[6] R.P. Anstee, A. Sali, *Small forbidden configurations IV,* Combinatorica 25 (2005): 503–518.

[7] C. Berge, Hypergraphs: Combinatorics of Finite Sets, North-Holland, Amsterdam (1989).

[8] N. Bousquet, A. Lagoutte, S. Thomassé, *The Erdős-Hajnal conjecture for paths and antipaths,* Journal of Combinatorial Theory, Series B, 113 (2015): 261–264.

[9] S. Buzalgo, R. Holzman, R. Pinchasi, *On s-intersecting curves and related problems,* Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry (2008): 79–84.

[10] M. Chudnovsky, *The Erdős-Hajnal conjecture – a survey,* Journal of Graph Theory 75 (2014): 178–190.

[11] M. Chudnovsky, A. Scott, P. Seymour, S. Spirkl, *Pure pairs. I. Trees and linear anticomplete pairs,* arXiv:1809.00919 preprint.

[12] P. Erdős, A. Hajnal, *On spanned subgraphs of graphs,* Graphentheorie und Ihre Anwendungen (Oberfoh, 1977)

[13] P. Erdős, A. Hajnal, *Ramsey type theorems,* Discrete Applied Mathematics 25 (1989): 37–52.

[14] J. Fox, J. Pach, *A Bipartite Analogue of Dilworth's Theorem for Multiple Partial Orders*, European Journal of Combinatorics 30 (2009): 1846–1853.

[15] J. Fox, J. Pach, *String graphs and incomparability graphs*, Advances in Mathematics 230 (2012): 1381–1401.

[16] J. Fox, J. Pach, C.D. Tóth, *Intersection patterns of curves*, Journal of the London Mathematical Society 83 (2011): 389–406.

[17] Z. Füredi, *The maximum number of unit distances in a convex n-gon*, Journal of Combinatorial Theory, Series A 55 (1990): 316–320.

[18] Z. Füredi, P. Hajnal, *Davenport-Schinzel theory of matrices*, Discrete Mathematics 103 (1992): 233–251.

[19] M.C. Golumbic, C.F. Goss, *Perfect elimination and chordal bipartite graphs*, Journal of Graph Theory 2 (1978): 155–163.

[20] A.J. Hoffman, A.W.J. Kolen, M. Sakarovitch, *Totally-balanced and greedy matrices*, SIAM Journal on Algebraic Discrete Methods 6 (1985): 721–730.

[21] B. Keszegh, D. Pálvölgyi, *An abstract approach to polychromatic coloring: shallow hitting sets in ABA-free hypergraphs and pseudohalfplanes*, Journal of Computational Geometry 10 (1) (2019): 1–26.

[22] B. Klinz, R. Rudolf, G.J. Woeginger, *Permuting matrices to avoid forbidden submatrices*, Discrete Applied Mathematics 60 (1995): 223–248.

[23] D. Korándi, G. Tardos, I. Tomon, C. Weidert, *On the Turán number of ordered forests*, Journal of Combinatorial Theory, Series A 165 (2019): 32–43.

[24] D. Korándi, I. Tomon, *Improved Ramsey-type results for comparability graphs*, Combinatorics, Probability and Computing, to appear.

[25] A. Liebenau, M. Pilipczuk, P. Seymour, S. Spirkl, *Caterpillars in Erdős-Hajnal*, Journal of Combinatorial Theory, Series B 136 (2019): 33–43.

[26] L. Lovász, *Graphs and set-systems*, in: Beiträge zur Graphentheorie, Teubner, Leipzig (1968): 99–106.

[27] L. Lovász, *Perfect graphs*, in: Selected Topics in Graph Theory, vol. 2, Academic Press, London (1983): 55–87.

[28] A. Lubiw, *Doubly lexical orderings of matrices*, SIAM Journal on Computing 18 (1987): 854–879.

[29] A. Marcus, G. Tardos, *Excluded permutation matrices and the Stanley-Wilf conjecture*, Journal of Combinatorial Theory, Series A 107 (2004): 153-160.

[30] J. Pach, G. Tardos, *Forbidden paths and cycles in ordered graphs and matrices*, Israel Journal of Mathematics 155 (2006): 359–380.

[31] J. Pach, G. Tóth, *How many ways can one draw a graph?*, Combinatorica 26 (2006): 559–576

[32] J. Pach, G. Tóth, *Comments on Fox News*, Geombinatorics 15 (2006): 150–154.

[33] V. Rödl, *On universality of graphs with uniformly distributed edges*, Discrete Mathematics 59 (1986): 125-134.

[34] A. Scott, P. Seymour, S. Spirkl, *Pure pairs. VII. Homogeneous submatrices in 0/1-matrices with a forbidden submatrix*, preprint.

[35] I. Tomon, *String graphs have the Erdős-Hajnal property*, arXiv:2002.10350 preprint.