A Refutation of the Pach-Tardos Conjecture for 0–1 Matrices

Seth Pettie*
University of Michigan

Gábor Tardos†
Rényi Institute

Abstract

The theory of forbidden 0–1 matrices generalizes Turán-style (bipartite) subgraph avoidance, Davenport-Schinzel theory, and Zarankiewicz-type problems, and has been influential in many areas, such as discrete and computational geometry, the analysis of self-adjusting data structures, and the development of the graph parameter *twin width*.

The foremost open problem in this area is to resolve the *Pach-Tardos conjecture* from 2005, which states that if a forbidden pattern $P \in \{0,1\}^{k\times l}$ is *acyclic*, meaning it is the bipartite incidence matrix of a forest, then $\mathrm{Ex}(P,n) = O(n\log^{C_P} n)$, where $\mathrm{Ex}(P,n)$ is the maximum number of 1s in a P-free $n\times n$ 0–1 matrix and C_P is a constant depending only on P. This conjecture has been confirmed on many small patterns, specifically all P with weight at most 5, and all but two with weight 6.

The main result of this paper is a clean refutation of the Pach-Tardos conjecture. Specifically, we prove that $\mathrm{Ex}(S_0,n), \mathrm{Ex}(S_1,n) \geq n2^{\Omega(\sqrt{\log n})}$, where S_0, S_1 are the outstanding weight-6 patterns.

$$S_0 = \begin{pmatrix} \bullet & & \bullet & \\ \bullet & & & \bullet \\ & \bullet & & \bullet \end{pmatrix}, \qquad S_1 = \begin{pmatrix} \bullet & & & \bullet \\ \bullet & & \bullet & \\ & \bullet & & \bullet \end{pmatrix}, \qquad P_t = \begin{pmatrix} \bullet & \overbrace{\bullet \quad & \quad & \bullet & \cdots & & \bullet}^{t+1 \text{ alternating 1s}} \\ \bullet & & \bullet & & & \bullet & \end{pmatrix}.$$

We also prove sharp bounds on the entire class of *alternating* patterns (P_t), specifically that for every $t \geq 2$, $\mathrm{Ex}(P_t,n) = \Theta(n(\log n/\log\log n)^t)$. This is the first proof of an asymptotically sharp bound that is $\omega(n\log n)$.

1 Introduction

The extremal theory of pattern-avoiding 0–1 matrices kicked off in the late 1980s when Mitchell [Mit92], Pach and Sharir [PS91], and Füredi [Für90] applied forbidden matrix arguments to problems in discrete and computational geometry. In the early days this theory was characterized [Mit92, FH92] as a *two dimensional* generalization of Davenport-Schinzel theory [SA95, Pet15a]. It can also be characterized as a generalization of Turán theory [Tur41, FS13] from unordered bipartite graphs to *ordered* bipartite graphs. Füredi and Hajnal [FH92] (see also Bienstock and Győri [BG91]) began the daunting project of classifying all forbidden patterns by their extremal function, a project to which many researchers have made important contributions over the years [Kla92, KV94, Tar05, PT06, Kes09, Cib09, Ful09, Gen09, Pet11a, Pet11b, Fox13, PS13, Pet15a, Pet15b, CK17, GT17, WP18, GKM+18, KTTW19, Gen19, FKMV20, GT20, MT22, GMN+23, KT23a, KT23b, JJMM24, CPY24, PT24]. Before proceeding let us define the terms.

1.1 Forbidden Patterns, 0–1 Matrices, Extremal Functions

A matrix $A \in \{0,1\}^{n\times m}$ *contains* a pattern $P \in \{0,1\}^{k\times l}$, written $P \prec A$, if it is possible to transform A into P by removing rows and columns from A, and flipping 1s to 0s. If $P \nprec A$ we say A is P-*free*. Let $\mathcal{P}$ be a *set* of forbidden patterns. The general *extremal function* is defined as follows.

$$\mathrm{Ex}(\mathcal{P},n,m) = \max\{\|A\|_1 \;\mid\; A \in \{0,1\}^{n\times m} \text{ and } \forall P \in \mathcal{P}. P \nprec A\},$$

*Supported by NSF Grant CCF-2221980. Email: pettie@umich.edu.
†Supported by the National Research, Development and Innovation Office projects K-132696 and SNN-135643 and by the ERC Advanced Grants "ERMiD" and "GeoScape." Email: tardos@renyi.hu.

where $\|A\|_1$, the *weight* of the matrix A, is the number of 1s in it. When there is a single forbidden pattern or A is square we use the short forms

$$\begin{aligned} \mathrm{Ex}(P,n,m) &= \mathrm{Ex}(\{P\},n,m), \\ \mathrm{Ex}(P,n) &= \mathrm{Ex}(P,n,n). \end{aligned}$$

$P \in \{0,1\}^{k\times l}$ can be regarded as the adjacency matrix of a bipartite graph with $k+l$ vertices, where the parts of the partition (rows and columns) are implicitly *ordered*. Define $G(P)$ to be the unordered bipartite graph corresponding to P. Turán's extremal function $\mathrm{Ex}_{\mathit{Tur}}(H,n)$ is defined to be the maximum number of edges in a simple n-vertex graph not containing H as a subgraph.

1.2 The Classification of Patterns We have a crude classification of forbidden subgraphs according to the asymptotic behavior of their Turán-extremal functions.

- If H is non-bipartite, then $\mathrm{Ex}_{\mathit{Tur}}(H,n) = \Theta(n^2)$.[1]
- If H is bipartite and contains a cycle, then $\mathrm{Ex}_{\mathit{Tur}}(H,n) = \Omega(n^{1+c_1})$ and $O(n^{1+c_2})$ for $0 < c_1 < c_2 < 1$.[2]
- If H is acyclic (a forest) then $\mathrm{Ex}_{\mathit{Tur}}(H,n) = \Theta(n)$.

Is there a similarly clean asymptotic classification for forbidden patterns in 0–1 matrices? In a very influential paper, Füredi and Hajnal [FH92] observed that (trivially) $\mathrm{Ex}(P,n) = \Omega(\mathrm{Ex}_{\mathit{Tur}}(G(P),2n))$ and that there were several examples when $\mathrm{Ex}(P,n) = \omega(\mathrm{Ex}_{\mathit{Tur}}(G(P),2n))$, e.g.,[3]

$$P_1 = \begin{pmatrix} \bullet & \bullet & \\ \bullet & & \bullet \end{pmatrix}, \qquad\qquad Q_3 = \begin{pmatrix} \bullet & & \bullet & \\ & \bullet & & \bullet \end{pmatrix}.$$

Both are acyclic matrices, so $\mathrm{Ex}_{\mathit{Tur}}(G(P_1),n) = \mathrm{Ex}_{\mathit{Tur}}(G(Q_3),n) = O(n)$, but $\mathrm{Ex}(P_1,n) = \Theta(n\log n)$ [BG91, Für90, FH92, Tar05] and $\mathrm{Ex}(Q_3,n) = \Theta(n\alpha(n))$ [FH92, HS86], where α is the inverse-Ackermann function. The pattern P_1 arises in an analysis of the Bentley-Ottman line-sweeping algoithm [PS91], bounding unit distances in convex n-gons [Für90], and bounding the total length of path compressions on arbitrary trees [Pet10]. The pattern Q_3 corresponds to order-3 (*ababa*-free) Davenport-Schinzel sequences, which have applications to lower envelopes [WS88, SA95] and corollaries of the dynamic optimality conjecture [CGJ+23, Pet08, CGK+15a, CPY24].

1.3 The Füredi-Hajnal and Pach-Tardos Conjectures Füredi and Hajnal made three conjectures concerning the relationship between Ex and $\mathrm{Ex}_{\mathit{Tur}}$.

Conjecture 1.1. (Füredi and Hajnal [FH92]) *If P is a permutation matrix (equivalently, $G(P)$ is a matching), then* $\mathrm{Ex}(P,n) = O(\mathrm{Ex}_{\mathit{Tur}}(G(P),n)) = O(n)$.

Conjecture 1.2. ([FH92]) *For any P,* $\mathrm{Ex}(P,n) = O(\mathrm{Ex}_{\mathit{Tur}}(G(P),n)\cdot\log n)$.

Perhaps doubting the validity of Conjecture 1.2 *in general*, they asked whether it held at least for acyclic patterns.

Conjecture 1.3. ([FH92]) *For any acyclic P,* $\mathrm{Ex}(P,n) = O(\mathrm{Ex}_{\mathit{Tur}}(G(P),n)\cdot\log n) = O(n\log n)$.

In 2004, Marcus and Tardos [MT04] proved Conjecture 1.1, which also proved the *Stanley-Wilf conjecture*, via a prior reduction of Klazar [Kla00]. This result inspired a line of research that led to the definition of the graph parameter *twin width* [GM14, BGK+21, BKTW22]. Although the leading constant in $\mathrm{Ex}(P,n) = O(n)$ for a k-permutation P depends only on k, it is *exponentially* larger than the corresponding leading constant of $\mathrm{Ex}_{\mathit{Tur}}(G(P),n) = O(n)$; see Fox [Fox13] and Cibulka and Kynčl [CK17].

[1] Erdős, Stone, and Simonovits [ES46, ES66], generalizing Turán's theorem [Tur41] bounded it more precisely as $\mathrm{Ex}_{\mathit{Tur}}(H,n) = (1-1/r+o(1))\binom{n}{2}$ if H has chromatic number $r+1$.

[2] Erdős and Simonovits conjectured that $\mathrm{Ex}(H,n) = \Theta(n^{1+\alpha})$ for some rational $\alpha \in \mathbb{Q}$; see [FS13].

[3] Following convention, we write patterns using bullets for 1s and blanks for 0s.

In 2005 Pach and Tardos [PT06] refuted Conjecture 1.2. They provided a matrix with $\|A\|_1 = \Theta(n^{4/3})$ that for each k, avoids a certain pattern D_{2k} for which $G(D_{2k}) = C_{2k}$ is a $2k$-cycle. Since $\mathrm{Ex}_{Tur}(C_{2k}, n) = O(n^{1+1/k})$ [BS74], this proved that the gap between $\mathrm{Ex}(P, n)$ and $\mathrm{Ex}_{Tur}(G(P), n)$ can be as large as $n^{1/3-\epsilon}$ for any $\epsilon > 0$. This result had no direct effect on Conjecture 1.3, but cast some doubt on its validity. Before Conjecture 1.3 was refuted they stated a more plausible version of it.

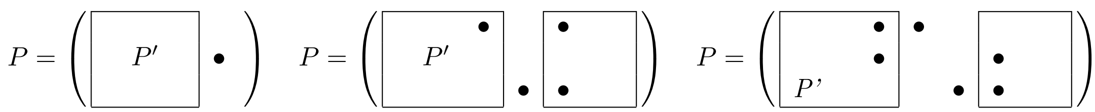
Conjecture 1.4. (Pach and Tardos [PT06]) *Let P be an acyclic 0–1 pattern.*

Weak Version. $\mathrm{Ex}(P, n) = O(n \log^{C_P} n)$*, for* some *constant* C_P*.*

Strong Version. $\mathrm{Ex}(P, n) = O(n \log^{\|P\|_1 - 3} n)$.

The rationale for the **Strong Version** is that all acyclic P with weight 3 are known to be linear [FH92], and in *some circumstances*, adding a row/column containing a single 1 only increases the extremal function by a $\log n$ factor. In particular, Pach and Tardos [PT06] proved the following three reductions for eliminating weight-1 columns. (In the diagrams, there are no constraints on the order of the rows.)

Lemma 1.1. (Pach and Tardos [PT06]) *Suppose P is obtained from P' (marked by boxes) by adding weight-1 columns in the following configurations.*[4]



(A) (B) (C)

Then $\mathrm{Ex}(P, n)$ *can be expressed in terms of* $\mathrm{Ex}(P', n)$ *as follows.*

(A) $\mathrm{Ex}(P, n) = O(\mathrm{Ex}(P', n) \log n)$.

(B) $\mathrm{Ex}(P, n) = O(\mathrm{Ex}(P', n) \log n)$.

(C) $\mathrm{Ex}(P, n) = O(\mathrm{Ex}(P', n) \log^2 n)$.

The Pach-Tardos reductions (Lemma 1.1) are sufficient to prove Conjecture 1.4 on all patterns with weight at most 5 and most of weight 6. For example, consider the class (P_t) of "alternating" patterns and R_0, R_1, R_2.

$$P_t = \begin{pmatrix} \bullet & \overbrace{\bullet & & \bullet & \cdots & & \bullet}^{t+1 \text{ alternating 1s}} \\ \bullet & & \bullet & & \cdots & \bullet & \end{pmatrix}, \qquad R_0 = \begin{pmatrix} \bullet & \bullet & \\ & & \bullet \\ \bullet & & \bullet \end{pmatrix},$$

$$R_1 = \begin{pmatrix} \bullet & \bullet & & \\ & & \bullet & \bullet \\ \bullet & & & \bullet \end{pmatrix}, \qquad R_2 = \begin{pmatrix} \bullet & \bullet & & \\ \bullet & & & \bullet \\ & & \bullet & \bullet \end{pmatrix}.$$

A t-fold application of Lemma 1.1(A) implies $\mathrm{Ex}(P_t, n) = O(n \log^t n)$, (B) implies $\mathrm{Ex}(R_0, n) = O(n \log^2 n)$, and (C) implies $\mathrm{Ex}(R_1, n) = O(n \log^3 n)$ and $\mathrm{Ex}(R_2, n) = O(n \log^2 n)$. However, there are two weight-6 patterns up to rotation/reflection that the Pach-Tardos reductions cannot simplify, namely S_0 and S_1.

$$S_0 = \begin{pmatrix} \bullet & & \bullet & \\ \bullet & & & \bullet \\ & \bullet & & \bullet \end{pmatrix}, \qquad S_1 = \begin{pmatrix} \bullet & & & \bullet \\ \bullet & & \bullet & \\ & \bullet & & \bullet \end{pmatrix}.$$

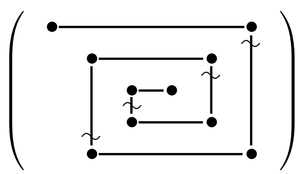
[4]Formally: (A) The last column of P has one 1. (B) Column j of P has one 1. There are rows i_0, i_1 such that $P(i_0, j) = P(i_0, j+1) = P(i_1, j-1) = P(i_1, j+1) = 1$. (C) Columns j and $j+1$ of P have one 1 each. There are rows i_0, i_1, i_2 such that $P(i_0, j-1) = P(i_0, j) = P(i_1, j+1) = P(i_1, j+2) = P(i_2, j-1) = P(i_2, j+2) = 1$.
Since P and its transpose have the same extremal function, reductions (A), (B), and (C) can also be applied to remove weight-1 rows. Strictly speaking, part (C) is implied by part (B).

1.4 Acyclic Patterns and the Status of Conjectures 1.3 and 1.4 In 2010 Pettie [Pet11a] refuted Füredi and Hajnal's Conjecture 1.3 by exhibiting an acyclic pattern X for which $\mathrm{Ex}(X, n) = \Omega(n \log n \log\log n)$.

$$X = \begin{pmatrix} & \bullet & & \bullet & \bullet \\ & & \bullet & & \bullet \\ & \bullet & & & \\ \bullet & & & & \bullet \end{pmatrix}.$$

Park and Shi [PS13] generalized this construction to a class (X_m) of acyclic patterns for which $\mathrm{Ex}(X_m, n) = \Omega(n \log n \log\log n \log\log\log n \cdots \log^{(m)} n)$. These results [Pet11a, PS13] did not cast any doubt on the Pach-Tardos conjecture (Conjecture 1.4), and even left open the possibility that the Füredi-Hajnal conjecture (Conjecture 1.3) was still *morally true*, e.g., if $\mathrm{Ex}(P, n) = O(n \log n \,\mathrm{poly}(\log\log n))$ for all acyclic P.

Essentially no progress has been made on expanding Pach and Tardos's repertoire of weight-1 column reduction rules (Lemma 1.1) in order to put more acyclic matrices in the $n\,\mathrm{poly}(\log n)$ class. However, in 2019 Korándi, Tardos, Tomon, and Weidert [KTTW19] developed a new technique for analyzing S_0, S_1 and similar matrices. They defined a pattern S to be *class-s degenerate* if it can be written $S = \begin{pmatrix} S' \\ S'' \end{pmatrix}$, where at most one column has a non-zero intersection with both S' and S'', and S', S'' are at most class-$(s-1)$ degenerate; any pattern with a single row is class-0 degenerate. Here is an example of a class-4 degenerate pattern. It is decomposed into individual rows by sequentially making horizontal cuts, each one cutting one vertical line segment joining 1s in the same column.[5]



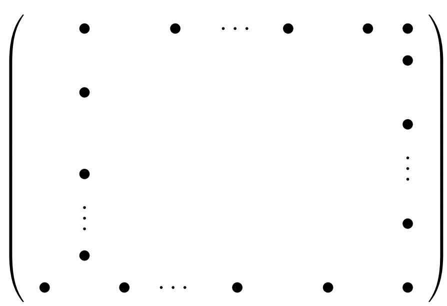
They proved that every class-s degenerate S has

$$\mathrm{Ex}(S, n) \le n \cdot 2^{O(\log^{1-\frac{1}{s+1}} n)} = n^{1+o(1)}.$$

As a consequence, $\mathrm{Ex}(S_0, n), \mathrm{Ex}(S_1, n) \le n2^{O(\log^{2/3} n)}$, and by being more careful with the analysis of S_0, they proved $\mathrm{Ex}(S_0, n) \le n2^{O(\sqrt{\log n})}$. The results of Korándi et al. [KTTW19] did not directly challenge the Pach-Tardos conjecture (Conjecture 1.4), and the authors characterized their results as taking a step towards *affirming* Conjecture 1.4.

Pettie and Tardos [PT24] introduced a class of matrices (A_t) such that A_t is B_t-free and $\|A_t\|_1 = \Theta(n(\log n / \log\log n)^t)$. The *box* pattern B_t is a $2t \times (2t+1)$ matrix, where the first and last rows form a reflection of P_{t+1} and the second and last columns form a rotation of P_t.

$B_t =$



Hence

$$\mathrm{Ex}(B_t, n) = \begin{cases} \Omega(n(\log n / \log\log n)^t) \\ O(n \log^{4t-3} n), \end{cases}$$

[5] We could also define degeneracy w.r.t. vertical cuts, i.e., S is class-s degenerate if $S = (S' \; S'')$, where S', S'' have at most one non-zero row in common and are at most class-$(s-1)$ degenerate. However, the Korándi et al. [KTTW19] method does not permit decomposing a pattern with *both* vertical and horizontal cuts.

where the upper bound follows from iterated application of Lemma 1.1(B). The Pettie-Tardos [PT24] lower bounds are the highest obtainable lower bounds that are *consistent* with the weak Pach-Tardos conjecture.[6]

1.5 Extensions and Variants of the Pach-Tardos Conjecture

Füredi, Jiang, Kostochka, Mubayi, and Verstraëte [FJK+21] studied forbidden patterns in ordered r-uniform hypergraphs. They made a conjecture extending Conjecture 1.4.

CONJECTURE 1.5. ([FJK+21, CONJECTURE B]) *Let F be any r-uniform forest with interval chromatic number r. Then the maximum number of edges in a vertex-ordered r-uniform hypergraph with no subgraph order-isomorphic to F is $O(n^{r-1}\log^{c(F)} n)$, for some constant $c(F)$.*

One can think of Conjecture 1.5 as a collection of separate conjectures for each $r \geq 2$. For $r = 2$ we get back an equivalent form of Conjecture 1.4; see [PT06, Theorem 2].

Shapira and Yuster [SY17] considered an extremal problem on augmented tournaments. A *tournament* is a complete graph with $\binom{n}{2}$ edges, each of which is assigned some direction. A t-augmented tournament has t extra directed edges, i.e., t pairs of vertices $\{u, v\}$ have both edges $(u, v), (v, u)$. Shapira and Yuster defined $t(n, H)$ to be the minimum number such that *any* n-vertex, $t(n, H)$-augmented tournament contains a subgraph isomorphic to the tournament H. They defined a notion of "tournament forest" and made an analogue of the Pach-Tardos conjecture.

CONJECTURE 1.6. ([SY17, CONJECTURE 1]) *For any tournament forest H there exists a constant c_H such that $t(n, H) = O(n \log^{c_H} n)$.*

Moreover, they proved that Conjecture 1.6 is equivalent to the Pach-Tardos conjecture.

THEOREM 1.1. (SHAPIRA AND YUSTER [SY17, THEOREM 1]) *Conjectures 1.4 and 1.6 are equivalent.*

1.6 New Results

Our main result is a *refutation* of the Pach-Tardos conjecture (Conjecture 1.4) in both its *weak* and *strong* forms. As stated above, the Pach-Tardos conjecture is equivalent to both the $r = 2$ case of the Füredi et al. conjecture (Conjecture 1.5) and the Shapira-Yuster conjecture (Conjecture 1.6) so both of the latter conjectures are also refuted. It is straightforward to modify the counterexample of the $r = 2$ case of Conjecture 1.5 (a graph being a 2-uniform hypergraph) to obtain counterexamples for any $r > 2$; see Appendix A.

Specifically, we prove that the two weight-6 patterns S_0, S_1 not subject to the Pach-Tardos reductions (Lemma 1.1) do not have $n\,\mathrm{poly}(\log n)$ extremal functions. Here and throughout the paper log stands for the binary logarithm.

THEOREM 1.2. $\mathrm{Ex}(S_0, n), \mathrm{Ex}(S_1, n) \geq n2^{\sqrt{\log n} - O(\log\log n)}$.

Theorem 1.2 matches Korándi et al.'s [KTTW19] upper bound $\mathrm{Ex}(S_0, n) \leq n2^{O(\sqrt{\log n})}$, up to the hidden constant in the exponent, which happens to be 4 in the upper bound rather than the 1 in the lower bound. We extend Theorem 1.2 in two directions. First, we show that the matrices constructed for the proof of this theorem avoid a large class of matrices beyond S_0 and S_1; see Theorem 2.2. Second, we modify the construction to increase its weight to $n2^{C\sqrt{\log n}}$ for any desired constant $C \geq 1$ and show that the matrices obtained still avoid some acyclic patterns; see Theorem 2.3.

We must admit that we *did not* specifically set out to disprove the Pach-Tardos conjecture. Our initial aim was simply to better understand variations on the construction of [PT24], and to understand simple, structured patterns like the (P_t) class. This effort was also very successful.

THEOREM 1.3. *For every $t \geq 2$, $\mathrm{Ex}(P_t, n) = \Theta(n(\log n/\log\log n)^t)$.*

[6](This is true in the sense that for any t [PT24] can achieve an $\Omega(n \log^t n)$ lower bound, but cannot obtain any $\Omega(n \log^{\omega(1)} n)$ lower bound.)

Both the upper and lower bounds of Theorem 1.3 are new. No lower bound better than $\mathrm{Ex}(P_t, n) \geq \mathrm{Ex}(P_1, n) = \Omega(n \log n)$ [FH92, BG91, Für90, Tar05] was previously known, and the best upper bound was $O(n \log^t n)$, which follows from iterated application of Lemma 1.1(A). Theorem 1.3 is notable in many ways. It is the *first* proof of an asymptotically sharp bound for any acyclic pattern with extremal function $\omega(n \log n)$; it demonstrates that the $(\log n/ \log\log n)^t$ density first seen in [PT24] is not contrived but a *natural* phenomenon, and it highlights an unexpected discontinuity between P_1 and $P_2, P_3, \ldots$

Although $2^{\sqrt{\log n}}$ and $(\log n/ \log\log n)^t$ look like they arise from quite different constructions, Theorems 1.2 and 1.3 use *essentially* the same 0–1 matrix construction for their lower bounds, but under different parameterizations.

1.7 Related and Unrelated Results

Unrelated Results. The function $2^{(\log n)^\delta}$ is a most fashionable function these days. Kelley and Meka [KM23] recently proved that Behrend's [Beh46] 1946 construction of 3-progression-free subsets of $[N] = \{1, 2, \ldots, N\}$ with size $N/2^{\Theta(\sqrt{\log N})}$ is roughly the best possible. Specifically, no 3-progression-free subset of $[N]$ has density $2^{-O(\log N)^{1/12}}$. This was later improved to $2^{-O(\log N)^{1/9}}$ by Bloom and Sisask [BS23]. Abboud, Fischer, Kelley, Lovett, and Meka [AFK+24] discovered a combinatorial boolean matrix multiplication algorithm running in $n^3/2^{\Omega(\log n)^{1/7}}$ time, improving a long line of $n^3/\operatorname{poly}(\log n)$-time algorithms.

Fine-grained Classification of Acyclic Patterns. Every existing analysis of an acyclic pattern P has placed its extremal function in the following five-echelon hierarchy. The first four echelons are "natural" inasmuch as there are lower bounds (see [HS86, FH92] and Theorem 1.2) proving that certain patterns in **Quasilinear**, **Polylog**, and **Near-linear** cannot be moved to a lower echelon.

Linear. $\mathrm{Ex}(P, n) = O(n)$.

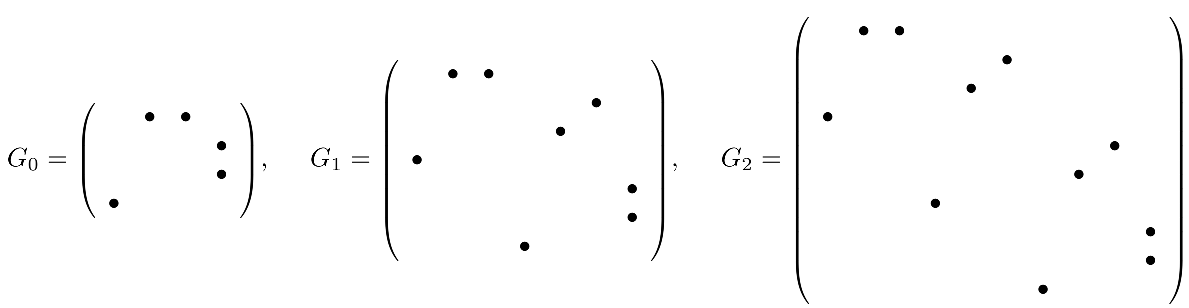
Quasilinear. $\mathrm{Ex}(P, n) = O(n 2^{(\alpha(n))^{C_P}})$, where $\alpha(n)$ is the inverse-Ackermann function.[7]

Polylog. $\mathrm{Ex}(P, n) = O(n \log^{C_P} n)$, for some $C_P > 0$.

Near-linear. $\mathrm{Ex}(P, n) = n 2^{O(\log^{1-\delta} n)}$, for some $\delta = \delta_P \in (0, 1)$.

Polynomial. $\mathrm{Ex}(P, n) = O(n^{1+C_P+o(1)})$, for some $C_P \in (0, 1)$.

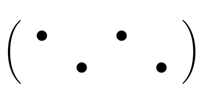
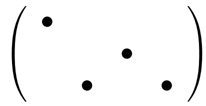
A lot of effort has been spent to understand the membership and boundaries of these classes. We know of some infinite classes of **Linear** matrices, such as permutations [MT04], double-permutations [Gen09], and monotone patterns [Kes09, Pet11c], and even have good bounds on the leading constant factors [Fox13, CK17, Gen15] for (double) permutations. Keszegh [Kes09] (see also [Gen09, Pet11a]) proved that the **Linear** class cannot be characterized by a finite set of minimally non-linear patterns. In particular, Lemma 1.1(A) and [Gen09] imply that every pattern in the infinite sequence (G_t) has $\mathrm{Ex}(G_t, n) = \Theta(n \log n)$ and there is an infinite sequence $(H_t), H_t \prec G_t$, of minimally non-linear patterns w.r.t. $\prec$ [Kes09, Gen09].[8]



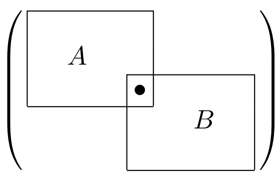
[7] There are more slowly growing functions in this class, e.g., $n\alpha(n)$ [HS86, FH92] or $n\alpha^2(n)$ [Pet11b, Pet15b].
[8] I.e., it is not known if G_t is itself minimally non-linear.

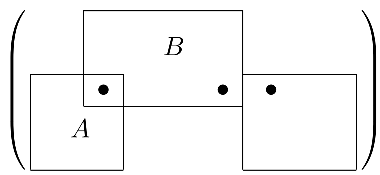
While the characterization of linear patterns seems to be elusive, Füredi, Kostochka, Mubayi and Verestrate [FKMV20] gave a simple characterization of linear *connected* patterns[9]

A pattern P is called *light* if it contains exactly one 1 per column. Light patterns are closely related to (generalized) Davenport-Schinzel sequences [SA95, Kla02]; they are *all* known to be in **Quasilinear** [Kla92, Kla02, Kes09]. More specifically, for any light P with two rows, $\mathrm{Ex}(P,n)$ is one of $\Theta(n)$, $\Theta(n\alpha(n))$, $\Theta(n2^{\alpha(n)})$, $\Theta(n\alpha(n)2^{\alpha(n)})$, or $n2^{(1+o(1))\alpha^t(n)/t!}$, $t \geq 2$, and if P has more rows, $\mathrm{Ex}(P,n)$ is upper bounded by one of these functions or $n(\alpha(n))^{(1+o(1))\alpha^t(n)/t!}$ [Pet15a, Pet15b, Pet11b, HS86, FH92, Niv10]. It is an open problem whether *light linear* patterns are themselves characterized by a finite set of forbidden patterns. The only known minimally non-linear ones (w.r.t. $\prec$ and rotation/reflection) are Q_3, Q_3', with $\mathrm{Ex}(\{Q_3,Q_3'\},n) = 2n\alpha(n) + O(n)$ [Niv10, Pet15a, Pet15b].

$Q_3 =$ , $Q_3' =$ .

Aside from Lemma 1.1's weight-1 column-reduction rules, there are a couple more methods to bound the extremal function of composite patterns.

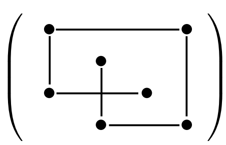
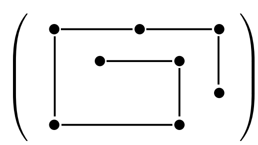
$A \oplus B =$ 

, $A \otimes B =$ 

.

Keszegh [Kes09] proved that $\mathrm{Ex}(A \oplus B, n) \leq \mathrm{Ex}(A,n) + \mathrm{Ex}(B,n)$[10] and Pettie [Pet11c], generalizing [Kes09], proved that if $\mathrm{Ex}(A,n)$ and $\mathrm{Ex}(B,n)$ are linear (respectively, near-linear) and B is *legal*[11] then $\mathrm{Ex}(A \otimes B, n)$ is linear (respectively, near-linear) as well.[12] These two composition rules put many more patterns in the **Linear** and **Polylog** classes [Kes09, Pet11c], and allow one to put "artificial" matrices in the class **Quasilinear** that are not light; see [PT24, Footnotes 3,4].

Every $P \in \{0,1\}^{k\times l}$ is dominated by the all-1 $k \times l$ matrix $K_{k,l}$. By the Kővári-Sós-Turán theorem, $\mathrm{Ex}(P,n) \leq \mathrm{Ex}(K_{k,l},n) = O(n^{2-1/\min\{k,l\}})$. Very recently Janzer, Janzer, Magner, and Methuku [JJMM24] proved that if P has at most t 1s per row (or per column), then $\mathrm{Ex}(P,n) = O(n^{2-1/t+o(1)})$, which confirmed a conjecture of Methuku and Tomon [MT22]. This upper bound applies to all patterns, but still provides the best known analysis for acyclic patterns not subject to Korándi et al.'s [KTTW19] method. For example, the *pretzel* and *spiral* patterns below have $\mathrm{Ex}(T_0,n), \mathrm{Ex}(T_1,n) = O(n^{3/2+o(1)})$.

$T_0 =$ , $T_1 =$ .

Little is known about the *gaps between* **Linear**, **Quasilinear**, and **Polylog**. For example, is there a pattern P with $\mathrm{Ex}(P,n) = \omega(n)$ but $o(n\alpha(n))$? or a P not in **Quasilinear** with $\mathrm{Ex}(P,n) = o(n\log n)$? Tardos [Tar05] proved that if we consider *pairs* of excluded matrices, $\mathrm{Ex}(\{P_1, P_1^{|}\},n) = \Theta(n\log n/\log\log n)$ while $\mathrm{Ex}(\{P_1, P_1^{\odot}\},n) = \Theta(n\log\log n)$, where $P_1^{|}$ is the reflection of P_1 across the y-axis and $P_1^{\odot}$ it its 180-degree rotation. Whether these and other extremal functions between quasilinear and $o(n\log n)$ can be realized by a *single* pattern is an open question.

[9] [FKMV20] is about connected vertex-ordered graphs but their results translate to the adjacency matrices of connected graphs too.

[10] $A \oplus B$ applies whenever A and B have 1s in their SE and NW corners, respectively.

[11] $A \otimes B$ applies when A contains two consecutive 1s in its top row, B has 1s in its SW and SE corners. A B is *legal* if it is either *non-ascending* or *non-descending*; see [Pet11c].

[12] Specifically, if $\mathrm{Ex}(A,n,m) \leq nf(n,m)$ and $\mathrm{Ex}(B,n,m) \leq ng(n,m)$ then $\mathrm{Ex}(A\otimes B,n,m) = O(nf(n,m)g(n,m))$.

Edge-ordered Graphs. As we noted earlier, 0–1 patterns can be regarded as bipartite graphs where the *vertex sets* on each side of the partition are ordered. Pach and Tardos [PT06] considered a vertex-ordered variant of Turán's problem where the vertex sets of the graph and pattern have a linear order. This theory is very similar to 0–1 matrix theory, with *interval chromatic number* 2 taking the role of *bipartiteness*. Gerbner et al. [GMN+23] introduced the analogous problem on edge-ordered graphs and forbidden patterns, where *order chromatic number* 2 takes the role of bipartiteness/interval chromatic number 2. The problem of characterizing linear edge-ordered patterns seems to be equally difficult in this setting; see [KT23b] characterizing *connected* edge-ordered graphs with a linear extremal function. Gerbner et al. [GMN+23] conjectured that the extremal function of edge-ordered acyclic patterns with order chromatic number 2 is $n^{1+o(1)}$. This conjecture was recently proved by Kucheriya and Tardos [KT23a], with a universal upper bound of $n2^{O(\sqrt{\log n})}$. They further conjectured that every such extremal function is actually bounded by $O(n \operatorname{poly}(\log n))$ *à la* Pach-Tardos. There is no known formal connection between the Gerbner et al./Kucheriya-Tardos conjectures and the Füredi-Hajnal/Pach-Tardos conjectures (Conjectures 1.3 and 1.4). For example, the refutations of the Füredi-Hajnal conjecture [Pet11a, PT24] had no analogues in the edge-ordered world, so one cannot expect an *automatic* way to transform Theorem 1.2 from vertex-ordered to edge-ordered graphs. In the reverse direction, [KT23a] has not led to a universal upper bound $\operatorname{Ex}(P, n) \le n2^{O(\sqrt{\log n})}$ on acyclic 0–1 patterns P. Nonetheless, Theorem 1.2 may cause one to doubt a universal $O(n \operatorname{poly}(\log n))$ upper bound for pattern graphs with order chromatic number 2.

1.8 Organization Section 2 introduces a new construction of 0–1 matrices with density (i.e., average number of 1s in a row) $\Theta(\log n/\log\log n)^t$ or $2^{\Theta(\sqrt{\log n})}$. It presents the full proof of Theorem 1.2 and its generalizations, as well as the lower bound half of Theorem 1.3. Section 3 presents the upper bound half of Theorem 1.3. We conclude with some open problems and conjectures in Section 4.

2 Lower Bounds

2.1 *The Good, the Bad, and the Ugly* 0–1 Matrix Construction We define a class of 0–1 matrices A w.r.t. integer parameters b and m. The rows are indexed by $[m] \times [m]^b$ and the columns by $[m]^b \times \{0,1\}^b$, both ordered lexicographically. Here $[m] = \{1, 2, \ldots, m\}$ is the set of the first m positive integers. If one desires a square matrix, then $m = 2^b$. We identify rows by pairs $\mathbf{r} = (s, r) \in [m] \times [m]^b$ and columns by pairs $\mathbf{c} = (c, i) \in [m]^b \times \{0,1\}^b$.

The matrix $A = A[b, m]$ is defined as follows[13]

$$A((s,r),(c,i)) = \begin{cases} 1 & \text{if } r = c + s \cdot i, \\ 0 & \text{otherwise.} \end{cases}$$

LEMMA 2.1. $\|A[b, m]\|_1 \ge 2^b m^b (m/(b+1) - 1)$, *i.e.,* $\Omega(m/b)$ *times the number of columns.*

Proof. Pick a column (c, i) uniformly at random. For some $s \in [m]$, there exists a row (s, r) such that $A((s,r),(c,i)) = 1$ iff, for all $u \in [b]$ with $i(u) = 1$, $r(u) = c(u) + s$ does not "overflow" the range $[m]$. In other words, the weight of column (c, i) is $m - \max_{u \in [b]\,:\, i(u)=1} c(u)$. We have

$$\mathbb{E}\left(\max_{u \in [b]\,:\, i(u)=1} c(u)\right) \le m - m/(\|i\| + 1) + 1 \le m - m/(b+1) + 1.$$

Thus, a column of A has at least $m/(b+1) - 1$ 1s on average. ∎

2.1.1 Matrices via Alternating Coordinate Offsets To construct P_t-free matrices we need a small modification to the $A[b, m]$ construction called $A_t = A_t[b, m]$. The row-set of A_t is the same as in $A[b, m]$, namely $[m] \times [m]^b$, but we only keep the following subset of the columns: $\{(c, i) \in [m]^b \times \{0,1\}^b \mid \|i\|_1 = t\}$. Thus, to form a square $A_t[b, m]$ one would set $m = \binom{b}{t}$. The rows and columns of A_t are also ordered lexicographically.

[13] By analogy to the plot of *The Good, the Bad, and the Ugly* (1966), the Ugly (column $\mathbf{c} = (c, i)$) knows the location of the graveyard (i) while the Good (row $\mathbf{r} = (s, r)$) knows the name on the grave (s).

Define the vectors $i^{\text{even}}, i^{\text{odd}} \in \{0,1\}^b$ for $i \in \{0,1\}^b$ as follows.

$$i^{\text{even}}(u) = i(u) \cdot \left(\left(1 + \sum_{v \geq u} i(v)\right) \bmod 2\right), \qquad i^{\text{odd}}(u) = i(u) \cdot \left(\left(\sum_{v \geq u} i(v)\right) \bmod 2\right).$$

Clearly, $i = i^{\text{even}} + i^{\text{odd}}$, with i^{odd} containing the last, 3rd last, 5th last 1s of i, and i^{even} containing the 2nd last, 4th last, 6th last 1s of i.

$$A_t((s,r),(c,i)) = \begin{cases} 1 & \text{if } r = c + s \cdot i^{\text{even}} + (m+1-s)i^{\text{odd}}, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 2.1. *It is possible to show that if one restricts $A = A[b,m]$ to columns (c,i) with $\|i\|_1 = t$, then A is P_{2t}-free. The reason for introducing A_t and the alternating offsets $s, m+1-s$ is to prove P_t-freeness.*

Lemma 2.2. *$\|A_t[b,m]\|_1 = \Omega(m^{b+1}\binom{b}{t}/(t2^t))$, i.e., $\Omega(m/(t2^t))$ times the number of columns.*

Proof. Pick a column (c,i) uniformly at random. For a fixed $s \in [m]$, there exists a row (s,r) with $A_t((s,r),(c,i)) = 1$ iff $c(u) + s \in [m]$ for all u with $i^{\text{even}}(u) = 1$ and $c(u) + m + 1 - s \in [m]$ for all u with $i^{\text{odd}}(u) = 1$. Thus, (c,i) will be incident to some row (s,r) for *every* $s \in [(1/2-\epsilon)m, (1/2+\epsilon)m]$ with probability at least $(1/2-\epsilon)^t$. Taking $\epsilon = 1/t$, the columns have $\Omega(m/(t2^t))$ 1s on average. □

We consider elements of $[m]^b$ as b-dimensional integer vectors and add/subtract them accordingly. When applied to vectors, '$<$' denotes lexicographic order. Note that lexicographic ordering makes these vectors into a linearly ordered group.

Lemma 2.3. (Simple Properties) *Let $A^* = A[b,m]$ or $A_t[b,m]$. Consider an occurrence in A^* of the following patterns.*

$$\begin{array}{c} \\ \mathbf{r}_0 \\ \mathbf{r}_1 \end{array} \overset{\mathbf{c}_0}{\begin{pmatrix} \bullet \\ \bullet \end{pmatrix}}, \qquad \mathbf{r}_0 \ \overset{\mathbf{c}_0 \ \ \mathbf{c}_1}{\begin{pmatrix} \bullet & \bullet \end{pmatrix}},$$

where $\mathbf{r}_0 = (s_0, r_0)$ and $\mathbf{c}_0 = (c_0, i_0)$, etc. Then on the left, $s_0 < s_1$, and on the right $c_0 < c_1$. If u is the position of the first non-zero of $c_1 - c_0$ then $i_0(u) = 1$. If $A^* = A[b,m]$, then $(c_1 - c_0)(u) = s_0$, whereas if $A^* = A_t[b,m]$ then $(c_1 - c_0)(u) = s_0$ if $i_0^{\text{even}}(u) = 1$ and $(c_1 - c_0)(u) = m + 1 - s_0$ if $i_0^{\text{odd}}(u) = 1$.

Proof. Observe that if $A((s,r),(c,i)) = 1$ then 'r' is uniquely determined by $s,(c,i)$ and 'i' is uniquely determined by $(s,r), c$.

Left. Rows are ordered primarily by their s component, so we must have $s_0 \leq s_1$. From the observation above, $s_0 = s_1$ is not possible.

Right. As above, we must have $c_0 \leq c_1$ and $i_0 = i_1$ is not possible, from the observation above. If $A^* = A[b,m]$ we have $c_1 - c_0 = s_0(i_0 - i_1)$. Here all coordinates of $i_0 - i_1$ are in $\{-1,0,1\}$. Now $i_0 \neq i_1$ implies that there is a first non-zero position u of $i_0 - i_1$ and $c_0 \leq c_1$ implies that $(i_0 - i_1)(u) = 1$ hence $(c_1 - c_0)(u) = s_0$.

If $A^* = A_t[b,m]$, then we have $c_1 - c_0 = s_0(i_0^{\text{even}} - i_1^{\text{even}}) + (m+1-s_0)(i_0^{\text{odd}} - i_1^{\text{odd}})$. Then $(i_0 - i_1)(u) = 1$ as before, and depending on whether $i_0^{\text{even}}(u) = 1$ or $i_0^{\text{odd}}(u) = 1$, we will see either $(c_1 - c_0)(u) = s_0$ or $m + 1 - s_0$. □

2.2 Refutation of the Pach-Tardos Conjecture

We now recall and prove Theorem 1.2.

Theorem 2.1. $\text{Ex}(S_0, n), \text{Ex}(S_1, n) \geq n2^{\sqrt{\log n} - O(\log\log n)}$.

Proof. We shall prove that for all b and m, $A = A[b,m]$ is $\{S_0, S_1\}$-free. The lower bound follows by setting $m = 2^b$, in which case A is an $n \times n$ matrix, $n = 2^{b^2+b}$ with $\|A\|_1 = \Omega(n(m/b))$. Here $b = \sqrt{\log n} - O(1)$.

Let us assume that S_0 is contained in A, embedded in the rows and columns as indicated below.

$$S_0 = \begin{array}{c} \\ \mathbf{r}_0 \\ \mathbf{r}_1 \\ \mathbf{r}_2 \end{array} \overset{\mathbf{c}_0 \quad \mathbf{c}_1 \quad \mathbf{c}_2 \quad \mathbf{c}_3}{\begin{pmatrix} \bullet & & \bullet & \\ \bullet & & & \bullet \\ & \bullet & & \bullet \end{pmatrix}},$$

where $\mathbf{r}_0 = (s_0, r_0)$ and $\mathbf{c}_0 = (c_0, i_0)$, etc.

Consider the differences $x = c_3 - c_0$, $y = c_2 - c_0$, and $z = c_3 - c_1$. We clearly have $y \le x$, $z \le x$, and $y + z \ge x$. Note that among the lexicographically ordered non-zero integer vectors, this implies that either (i) x and y agree on the position and value of their first non-zero coordinate, (ii) x and z agree on the position and value of their first non-zero coordinate, or (iii) the position of the first non-zero coordinate is the same for x, y and z and its value is strictly smaller for y and z than for x. Lemma 2.3, applied separately to the three rows, tells us that the first non-zero coordinates of x, y, and z are s_1, s_0 and s_2, respectively. Applying Lemma 2.3 to columns $\mathbf{c}_0$, and $\mathbf{c}_3$, we obtain that $s_0 < s_1 < s_2$, making cases (i), (ii), and (iii) impossible.

Note that S_1 is obtained from S_0 by swapping the first two rows. Without changing the definition of x, y, z, we still have $y, z \le x$ and $y + z \ge x$. Lemma 2.3 implies the first non-zero coordinates in x, y, z are s_0, s_1, s_2, respectively, and that $s_0 < s_1, s_2$. Once again, cases (i), (ii), and (iii) are all impossible. ∎

2.3 Extensions of Theorem 1.2 In this section we extend Theorem 1.2's lower bound on $\mathrm{Ex}(S_0, n)$ in two directions. In Section 2.3.1 we show that the argument of Theorem 1.2 can be applied to any *covering pattern*, a class whose simplest members are S_0 and S_1. In Section 2.3.2 we show that there is a class of S_0-like patterns whose extremal functions can reach $n2^{C\sqrt{\log n}}$ for any desired constant C.

2.3.1 Covering Patterns Definition 2.1 specifies the class of *covering patterns*.

DEFINITION 2.1. *Let $M \in \{0,1\}^{\alpha\times\beta}$ be an acyclic pattern with rows indexed with $0 \le i < \alpha$ and columns $0 \le j < \beta$. M is a* covering pattern *if there is a distinguished row k^* satisfying the following properties.*

1. *$M(k^*, 0) = M(k^*, \beta - 1) = 1$, i.e., row k^* has 1s in the first and last columns.*
2. *Let J be the set of row indices, excluding k^*, with at least two 1s. J contains at most one element $j_0 < k^*$. If such an element $j_0 \in J$ exists, then some column contains two 1s in rows $[j_0, k^*]$. For any element $j \in J$ with $j > k^*$ some column contains two 1s in rows $[k^*, j]$.*
3. *For $l \in J$ define $\mathsf{first}(l), \mathsf{last}(l)$ to be the column indices of the first and last 1s in row l. Then the* real *intervals $[\mathsf{first}(l), \mathsf{last}(l)]$ cover the entire range, i.e., we have $\bigcup_{l\in J}[\mathsf{first}(l), \mathsf{last}(l)] = [0, \beta - 1]$.*

S_0 and S_1 satisfy Definition 2.1, as do S_2, S_3. Note that S_0, S_2 have a $j_0 < k^*$ whereas S_1, S_3 do not.

$$S_2 = \begin{pmatrix} & \bullet & & & & \bullet & \\ & & & \bullet & & & \\ \bullet & & & \bullet & & & \bullet \\ \bullet & & \bullet & & & & \\ & & & & \bullet & & \bullet \end{pmatrix}, \qquad S_3 = \begin{pmatrix} \bullet & & & & & & & \bullet \\ & & & \bullet & & & & \\ & & & \bullet & & & \bullet & \\ & & & & & \bullet & & \bullet \\ & \bullet & & & \bullet & & & \\ \bullet & & \bullet & & & & & \end{pmatrix}.$$

THEOREM 2.2. *For every covering pattern M satisfying Definition 2.1, $A[b, m]$ is M-free, hence $\mathrm{Ex}(M, n) \ge n2^{\sqrt{\log n} - O(\log\log n)}$.*

Proof. Suppose, towards obtaining a contradiction, that $A = A[b, m]$ contains M, with row l of an instance of M labeled (s_l, r_l) and column j labeled (c_j, i_j). Definition 2.1(2) implies that all members of $\{s_l \mid l \in J\}$ are different from s_{k^*}, and all but at most one are strictly greater than s_{k^*}.

Define $x = c_{\beta-1} - c_0$, and for $l \in J$, $y_l = c_{\mathsf{last}(l)} - c_{\mathsf{first}(l)}$. Suppose the position of the first non-zero in x is u. By Lemma 2.3 applied to line k^* we have $x(u) = s_{k^*}$. Moreover, Lemma 2.3 implies the first non-zero in the vector y_l is s_l. We have $y_l \le x$ for all $l \in J$ hence $y_l(u) = 0$ for all $l > k^*$. (If $l < k^*$, it is possible that $y_l(u) = s_l < s_{k^*}$.) The covering property Definition 2.1(3) implies that $\sum_{l\in J} y_l \ge x$, but this cannot be satisfied since $\sum_{l\in J} y_l(u) < x(u) = s_{k^*}$. We conclude that A is M-free and by Lemma 2.1, $\mathrm{Ex}(M, n) \ge n2^{\sqrt{\log n} - O(\log\log n)}$. ∎

2.3.2 A Hierarchy of Patterns $S_0^{(t)}$ We prove that for every $C \geq 1$, there exists an acyclic pattern $S_0^{(t)}$ such that $\mathrm{Ex}(S_0^{(t)}, n) \geq n2^{C\sqrt{\log n}}$. The 0–1 matrix construction uses a small generalization of Behrend's arithmetic progression-free sets.

LEMMA 2.4. (CF. BEHREND [BEH46]) *For any $h \geq 2$, there exists a subset $S \subset [N]$ with $|S| = N/2^{O(\sqrt{\log h \log N})}$ such that there are no non-trivial solutions to $\alpha s_0 + \beta s_1 + \gamma s_2 = 0$, with $s_0, s_1, s_2 \in S$ and integers $-h \leq \alpha, \beta, \gamma \leq h$.*[14]

Proof. Let $V \subset \{0, \ldots, d-1\}^D$ be a subset of vectors with $|V| \geq d^{D-2}/D$ having a common ℓ_2-norm. Let us obtain S from V by prefixing each vector in V with a '1' and then interpreting them as $(D+1)$-digit integers in base $2hd$. In formula we have

$$S = \left\{ (2hd)^D + \sum_{i=0}^{D-1} v_i (2hd)^i \;\middle|\; (v_{D-1}, \ldots, v_1, v_0) \in V \right\}.$$

We set $d = \left\lfloor 2^{\sqrt{\log h \log(N/2)} - \log(2h)} \right\rfloor$ and $D = \left\lfloor \sqrt{\log(N/2)/\log h} \right\rfloor$, so for all $s \in S$ we have $s < 2(2hd)^D \leq N$. Expressed in terms of h, N, $|V| = |S| \geq d^{D-2}/D = N/2^{O(\sqrt{\log h \log N})}$.

Now suppose there is a solution to $\alpha s_0 + \beta s_1 + \gamma s_2 = 0$ with integers $-h \leq \alpha, \beta, \gamma \leq h$. We need to show this is a trivial solution. With a slight abuse of notation, we will identify s_0, s_1, s_2 with the sequence of $D+1$ digits in their base $2hd$ expression. Note that these sequences are all obtained from elements in V by putting an extra 1 in front of them, so they have the same ℓ_2-norms, say r. By symmetry, we may assume that $\gamma \leq 0 \leq \alpha, \beta$ and consider the equality $\alpha s_0 + \beta s_1 = |\gamma| s_2$ that holds on the level of integers. Due to the base being much larger than any of the digits, it must also hold for the corresponding sequence of digits. The first digit of $\alpha s_0 + \beta s_1$ is $\alpha + \beta$, while the first digit of $|\gamma| s_2$ is $|\gamma|$, so we must have $\alpha + \beta = |\gamma|$, $\alpha + \beta + \gamma = 0$. From $\|s_0\|_2 = \|s_1\|_2 = \|s_2\|_2 = r$ we obtain $\|(|\gamma| s_2)\|_2 = |\gamma| r$ but if $s_0 \neq s_1$ and α, β are positive, then $\|\alpha s_0 + \beta s_1\|_2 < (\alpha + \beta) r = |\gamma| r$, a contradiction. So if $\alpha, \beta > 0$, we have $s_0 = s_1$, and from $(\alpha + \beta) s_0 = |\gamma| s_2$, we also have $s_0 = s_1 = s_2$ and the solution is trivial. If one or both of α or β is zero, we similarly get that $\alpha s_0 + \beta s_1 + \gamma s_2 = 0$ constitutes a trivial solution. ◻

THEOREM 2.3. *For each $t \geq 2$, $\mathrm{Ex}(S_0^{(t)}, n) \geq n2^{(1-o(1))\sqrt{\log t \log n}}$, where $S_0^{(t)}$ is defined to be*

$$S_0^{(t)} = \begin{array}{c} \\ \mathbf{r}_0 \\ \mathbf{r}_1 \\ \mathbf{r}_2 \end{array} \!\!\begin{array}{c} \begin{array}{ccccccc} \mathbf{c}_0 & \mathbf{c}_1 & \mathbf{c}_2 & & \mathbf{c}_{2t-3} & \mathbf{c}_{2t-2} & \mathbf{c}_{2t-1} \end{array} \\ \left(\begin{array}{ccccccc} \bullet & & \bullet & & & \bullet & \\ \bullet & & & \cdots & & & \bullet \\ & \bullet & & & \bullet & & \bullet \end{array} \right) \end{array}.$$

Note that $S_0 = S_0^{(2)}$.

Proof. We modify the construction of $A = A[b, m]$ as follows. The rows and columns are identified with $S \times [m]^b$ and $[m]^b \times \{0, \ldots, t-1\}^b$, respectively, where $S \subset [m/b]$ is the set from Lemma 2.4 with size $m^{1-o(1)}$ avoiding solutions to $\alpha s_0 + \beta s_1 + \gamma s_2 = 0$ with integers $-t+1 \leq \alpha, \beta, \gamma \leq t-1$. As usual $A((s, r), (c, i)) = 1$ iff $r = c + si$. We make A square by choosing m such that $t^b = |S|$, so $m = t^{(1+o(1))b}$, $n = m^b t^b = t^{(1+o(1))b^2}$, and $b = (1 - o(1))\sqrt{\log n / \log t}$. There are $\Theta(|S|)$ 1s per column, on average, so $\|A\|_1 = \Omega(n|S|) = \Omega(n2^{(1-o(1))\sqrt{\log t \log n}})$.

We now argue that A is $S_0^{(t)}$-free. Assume, for a contradiction, that A contains $S_0^{(t)}$ in rows $\mathbf{r}_0 = (s_0, r_0), \mathbf{r}_1 = (s_1, r_1), \mathbf{r}_2 = (s_2, r_2)$ and columns $\mathbf{c}_0 = (c_0, i_0), \ldots, \mathbf{c}_{2t-1} = (c_{2t-1}, i_{2t-1})$. Define x, y_k, z_k as

$$\begin{aligned} x &= c_{2t-1} - c_0, & & \\ y_k &= c_{2k} - c_0, & & \text{for } 1 \leq k \leq t-1, \\ z_k &= c_{2t-1} - c_{2k-1}, & & \text{for } 1 \leq k \leq t-1. \end{aligned}$$

Observe that for all k, $y_k \leq x$ and $z_k \leq x$ but

$$y_k + z_k \geq x \geq y_k + z_{k+1}, \tag{2.1}$$

[14] A solution is trivial if it exists for any non-empty S, that is, $\alpha + \beta + \gamma = \alpha\beta(s_0 - s_1) = \beta\gamma(s_1 - s_2) = \gamma\alpha(s_2 - s_0) = 0$.

where the last inequality holds for $k < t-1$.

Note that the following analogue of Lemma 2.3 holds here: (i) if a column of A contains 1s in the distinct rows (s, r) and (s', r'), then $s \neq s'$ and (ii) if the row (s, r) of A contains 1s in the distinct columns (c, i) and (c', i'), then $c \neq c'$ and all coordinates of $c - c'$ are of the form js with $-t+1 \leq j \leq t-1$. In particular, (i) implies $s_0 < s_1 < s_2$.

Let u be the position of the first non-zero coordinate of x. By (ii) above and $c_0 \leq c_{2t-1}$ we have $x(u) = js_1$ with $1 \leq j \leq t-1$. We use (ii) again to define j_k, j'_k such that

$$y_k(u) = j_k s_0,$$
$$z_k(u) = j'_k s_2.$$

Note that y_k, z_k cannot have any non-zeros preceding coordinate u and must have a non-negative value at position u. Thus $0 \leq j_k \leq t-1$ and for (2.1) to be satisfied, $1 \leq j'_k < j-1$ since $s_2 > s_1$. (This is already a contradiction when $t = 2$.) We can write Eq. (2.1) as

$$j_k s_0 + j'_k s_2 \geq j s_1 \geq j_k s_0 + j'_{k+1} s_2, \tag{2.2}$$

Hence $t-1 \geq j > j'_1 \geq j'_2 \geq \cdots \geq j'_{t-1} \geq 1$. By the pigeonhole principle there must exist $j'_k = j'_{k+1}$, but then (2.2) holds with equality, meaning S supports a non-trivial solution to $\alpha s_0 + \beta s_1 + \gamma s_2 = 0$ with $\alpha = j_k, \beta = -j, \gamma = j'_k$, a contradiction. □

2.4 Polylogarithmic Lower Bounds on Alternating Matrices In this section we prove the lower bound half of Theorem 1.3.[15]

THEOREM 2.4. *For $t \geq 1$, $\mathrm{Ex}(P_t, n) = \Omega(n(\log n/\log\log n)^t)$.*

Proof. We prove that $A_t[b, m]$ is P_t-free. The lower bound follows by setting $m = \binom{b}{t}$, in which case A_t is an $n \times n$ 0–1 matrix, $n = m^{b+1} = \binom{b}{t}^{b+1}$, with $\Omega(m) = \Omega((\log n/\log\log n)^t)$ 1s per column, on average.

Suppose that A_t were not P_t-free. Label the rows and columns of an occurrence of P_t in A_t as follows, where $\mathbf{r}_0 = (s_0, r_0), \mathbf{c}_0 = (c_0, i_0)$, etc.

$$P_t = \begin{array}{c} \\ \mathbf{r}_0 \\ \mathbf{r}_1 \end{array} \begin{array}{c} \begin{array}{cccccccc} \mathbf{c}_0 & \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & & & \mathbf{c}_{t+1} \end{array} \\ \left(\begin{array}{cccccccc} \bullet & \bullet & & \bullet & \cdots & & \bullet \\ \bullet & & \bullet & & & \bullet & \end{array}\right) \end{array}.$$

Define $u_j \in [b]$ to be the position of the first non-zero of $c_j - c_0$. Note that $\mathbf{c}_j$ has its 1 in row $\mathbf{r}_{(j+1) \bmod 2}$. By Lemma 2.3, either $i_0^{\mathrm{even}}(u) = 1$ and $(c_j - c_0)(u_j) = s_{(j+1) \bmod 2}$ or $i_0^{\mathrm{odd}}(u) = 1$ and $(c_j - c_0)(u_j) = m+1-s_{(j+1) \bmod 2}$. In either case $(c_j - c_0)(u_j) > 0$, so from the ordering $c_0 \leq c_1 \leq \cdots \leq c_{t+1}$ we can conclude that

$$u_1 \geq u_2 \geq \cdots \geq u_{t+1}.$$

Since $\|i_0\|_1 = t$, by the pigeonhole principle there must exist a j with $u_j = u_{j+1}$. In fact, there must exist such a j that also satisfies j odd and $i_0^{\mathrm{odd}}(u_j) = 1$ or j even and $i_0^{\mathrm{even}}(u_j) = 1$.[16] If j is odd then

$$c_j - c_0 = (0, \ldots, 0, m+1-s_0, \ldots)$$
$$c_{j+1} - c_0 = (0, \ldots, 0, m+1-s_1, \ldots),$$

which contradicts the order $c_j \leq c_{j+1}$ as $s_0 < s_1$ by Lemma 2.3. If j is even then

$$c_j - c_0 = (0, \ldots, 0, s_1, \ldots)$$
$$c_{j+1} - c_0 = (0, \ldots, 0, s_0, \ldots),$$

which also contradicts the order $c_j \leq c_{j+1}$. Hence A_t is P_t-free. □

---

[15] In the context of Theorems 1.3, 2.4 and 3.1, t is fixed, and the constants hidden by O, Ω depend on t. This is in contrast to Lemma 2.2, which did not assume t is constant; its constants hidden by asymptotic notation are independent of t.

[16] Note that $u_1 = u_2$ is the last 1 of i_0, or $u_2 = u_3$ is the 2nd last 1 of i_0, etc., all of which satisfy the parity criterion.

3 Upper Bounds

Theorem 3.1 covers the upper bound half of Theorem 1.3. Note the condition $t \geq 2$ cannot be strengthened, as $\text{Ex}(P_1, n) = \Theta(n \log n)$ [FH92, BG91, Tar05], not $O(n \log n / \log\log n)$.

Theorem 3.1. *For $t \geq 2$, $\text{Ex}(P_t, n) = O(n(\log n/\log\log n)^t)$.*

Let A be an $n \times n$, P_t-free matrix maximizing $\|A\|_1$. For each $A(r,c) = 1$ we identify a number of *landmark* column indices.

Definition 3.1. (Landmark Columns) *With respect to some $A(r,c) = 1$, the following column indices obey the following order whenever they exist.*

$$c < F \leq a_1 < b_1 \leq a_2 < b_2 \leq \cdots \leq a_{t-1} < b_{t-1} \leq L.$$

- *$A(r,F) = A(r,L) = 1$ are the first and last 1s in row r following column c.*
- *$A(r,a_{t-1}) = A(r,b_{t-1}) = 1$ are consecutive 1s in row r such that $F \leq a_{t-1} < b_{t-1} \leq L$ and $b_{t-1} - a_{t-1}$ is maximum. In general, $A(r,a_j) = A(r,b_j) = 1$ are consecutive 1s in row r such that $F \leq a_j < b_j \leq a_{j+1}$ and $b_j - a_j$ is maximum. (Break ties in a consistent way, say taking the first pair of consecutive 1s of maximal distance.)*
- *If $A(r,c)$ is one of the last two 1s in row r, then F and L are not distinct indices and $a_1, b_1, \ldots, a_{t-1}, b_{t-1}$ do not exist. Whenever $F = a_j$, the indices $a_1, b_1, \ldots, a_{j-1}, b_{j-1}$ do not exist.*

We assign three *signatures* to every 1 in A that is not one of the last two 1s in its row.

Definition 3.2. (Signatures) *Let $\zeta = \zeta(n) \geq t$ be a parameter (to be optimized later) and $\epsilon \stackrel{\text{def}}{=} \frac{1}{6(\zeta+1)(t+2)}$. Each $A(r,c) = 1$ is assigned three* signatures *$\text{sig}_0(r,c), \text{sig}_1(r,c)$, and $\text{sig}_2(r,c)$. Any piece of a signature that depends on undefined values (e.g., if a_1, b_1 do not exist) is undefined.*

$\text{sig}_0(r,c)$ **:** *consists of the vector $(\lfloor \log_\zeta(F-c) \rfloor, \lfloor \log_\zeta(b_1 - a_1) \rfloor, \ldots, \lfloor \log_\zeta(b_{t-1} - a_{t-1}) \rfloor)$.*

$\text{sig}_1(r,c)$ **:** *consists of two parts, a vector $(\lfloor \log_{1+\epsilon}(a_1 - c) \rfloor, \ldots, \lfloor \log_{1+\epsilon}(a_{t-1} - c) \rfloor)$, and the position of each $\lfloor \log_{1+\epsilon}(b_j - c) \rfloor$ relative to the elements of this vector: larger, equal, or smaller.*

$\text{sig}_2(r,c)$ **:** *consists of three parts, a vector*

$$\left(\lfloor \log_{1+\epsilon}(b_1 - c) \rfloor, \ \lfloor \log_{1+\epsilon}(b_2 - c) \rfloor, \ \ldots, \ \lfloor \log_{1+\epsilon}(b_{t-1} - c) \rfloor \right),$$

the position of each $\lfloor \log_{1+\epsilon}(a_j - c) \rfloor$ relative to the elements of this vector (as in sig_1), and a vector

$$\left(\min\left\{ \left\lfloor \frac{a_2 - b_1}{(b_1 - c)/2} \right\rfloor, 3t \right\}, \ \ldots, \ \min\left\{ \left\lfloor \frac{a_{t-1} - b_{t-2}}{(b_{t-2} - c)/2} \right\rfloor, 3t \right\}, \ \min\left\{ \left\lfloor \frac{L - b_{t-1}}{(b_{t-1} - c)/2} \right\rfloor, 3t \right\} \right).$$

3.1 Structure of the Proof

We will eventually prove that after the following 4-step *marking* procedure, there will be no unmarked 1s in any P_t-free matrix A.

Step 1. In each row, mark the last two 1s and the last 1 of each sig_0-type.

Step 2. In each column, mark the last unmarked 1 of each sig_1-type that satisfies Inequality (I.1), defined in Lemma 3.1.

Step 3. In each row, mark the last t unmarked 1s of each sig_2-type.

Step 4. In each column, mark the last t unmarked 1s of each sig_2-type.

The number of distinct sig_0 signatures is $O(\log_\zeta^t n)$ while the number of sig_1 and sig_2 signatures is $O(\log_{1+\epsilon}^{t-1} n) = O((\zeta \log n)^{t-1})$, so the total number of marked 1s is $O(n(\log_\zeta^t n + (\zeta \log n)^{t-1}))$. (Remember that $t = O(1)$ is constant.) We choose $\zeta = (\log n)^{1/t}$ to roughly balance these contributions, and conclude that $\|A\|_1 = O(n(\log n/\log\log n)^t)$.

3.2 The Proof Lemmas 3.1 to 3.4 will be used to prove that no unmarked 1s in A remain after Steps 1–4.

LEMMA 3.1. *Suppose that $A(r, c^*) = 1$, having landmarks $(F, a_1, b_1, \ldots, L)$, is* not *the last 1 in its row with its* sig_0*-type. Then for* at least one *of the following t inequalities, the relevant landmarks exist and the inequality is satisfied.*

$$\text{(I.1)} \qquad a_1 - F > \tfrac{1}{3\zeta}(F - c^*),$$
$$\text{(I.2)} \qquad a_2 - b_1 > \tfrac{1}{3\zeta}(b_1 - c^*),$$
$$\vdots$$
$$\text{(I.}t-1) \qquad a_{t-1} - b_{t-2} > \tfrac{1}{3\zeta}(b_{t-2} - c^*),$$
$$\text{(I.}t) \qquad L - b_{t-1} > \tfrac{1}{3\zeta}(b_{t-1} - c^*).$$

Proof. Suppose that for $c' > c^*$, $A(r, c') = 1$ has the same sig_0-type as $A(r, c^*)$, with landmarks $(F', a_1', b_1', \ldots, a_{t-1}', b_{t-1}', L')$. Then c' lies in one of the following intervals.

$$[F, a_1), [a_1, a_2), \ldots, [a_{j-1}, a_j), \ldots, [a_{t-1}, L).$$

Case 1: $c' \in [F, a_1)$**.** Clearly $A(r, c^*)$ and $A(r, c')$ share a suffix of the landmarks, specifically $(a_1, b_1, \ldots, a_{t-1}, b_{t-1}, L) = (a_1', b_1', \ldots, a_{t-1}', b_{t-1}', L')$; only $F' \in (F, a_1]$ is different. Since $\lfloor \log_\zeta(F - c^*) \rfloor = \lfloor \log_\zeta(F' - c') \rfloor$, it must be that $a_1 - F \geq F' - c' > \frac{1}{\zeta}(F - c^*) > \frac{1}{3\zeta}(F - c^*)$.

Case 2: $c' = a_j$**.** Since there are no 1s in row r and the column interval (a_j, b_j), $F' = b_j$ and the landmark vectors agree on the suffixes $(a_{j+1}, b_{j+1}, \ldots, L) = (a_{j+1}', b_{j+1}', \ldots, L')$. We prove that at least one of Inequalities (I.1)–(I.$j+1$) is satisfied. If (I.1)–(I.j) are not satisfied, then

$$\begin{aligned}
b_j - c^* &= (b_j - a_j) + (a_j - c^*) \\
&\leq (b_j - a_j) + (1 + \tfrac{1}{3\zeta})(b_{j-1} - c^*) \\
&\quad \cdots \\
&\leq (b_j - a_j) + (1 + \tfrac{1}{3\zeta})(b_{j-1} - a_{j-1}) + \cdots + (1 + \tfrac{1}{3\zeta})^{j-1}(b_1 - a_1) + (1 + \tfrac{1}{3\zeta})^j (F - c) \\
&< (1 + \tfrac{1}{3\zeta})^t [(b_j - a_j) + \cdots + (b_1 - a_1) + (F - c^*)].
\end{aligned} \tag{3.3}$$

According to the common sig_0 type and the fact that $(a_j, b_j) = (c', F')$ we have

$$\lfloor \log_\zeta(b_j' - a_j') \rfloor + \cdots + \lfloor \log_\zeta(b_1' - a_1') \rfloor = \lfloor \log_\zeta(b_j - a_j) \rfloor + \cdots + \lfloor \log_\zeta(b_1 - a_1) \rfloor, \tag{3.4}$$
$$\lfloor \log_\zeta(b_j' - a_j') \rfloor = \lfloor \log_\zeta(b_j - a_j) \rfloor = \lfloor \log_\zeta(F' - c') \rfloor = \lfloor \log_\zeta(F - c^*) \rfloor. \tag{3.5}$$

Since all the landmarks $F', a_1', \ldots, b_j'$ lie in the range $[b_j, a_{j+1}]$, Eqs. (3.3) to (3.5) imply that

$$\begin{aligned}
2(a_{j+1} - b_j) &\geq \tfrac{1}{\zeta}[(b_j - a_j) + \cdots + (b_1 - a_1) + (F - c^*)] \\
&> \tfrac{1}{\zeta}(1 + \tfrac{1}{3\zeta})^{-t}(b_j - c^*),
\end{aligned}$$

and since $\zeta \geq t$, we have $(1 + \frac{1}{3\zeta})^{-t} > e^{-1/3} > 2/3$. Thus,

$$a_{j+1} - b_j > \tfrac{1}{3\zeta}(b_j - c^*),$$

i.e., Inequality (I.$j+1$) is satisfied.

Case 3: $c' \in [b_j, a_{j+1})$**.** (Here $a_t \stackrel{\text{def}}{=} L$.) The argument is identical to Case 2, except that since c' is also in $[b_j, a_{j+1})$, we can substitute for Eqs. (3.4) and (3.5) the following inequality.

$$\begin{aligned}
&\lfloor \log_\zeta(b_j' - a_j') \rfloor + \cdots + \lfloor \log_\zeta(b_1' - a_1') \rfloor + \lfloor \log_\zeta(F' - c') \rfloor \\
&= \lfloor \log_\zeta(b_j - a_j) \rfloor + \cdots + \lfloor \log_\zeta(b_1 - a_1) \rfloor + \lfloor \log_\zeta(F - c) \rfloor.
\end{aligned}$$

Thus, $a_{j+1} - b_j > \frac{1}{\zeta}(1 + \frac{1}{3\zeta})^{-t}(b_j - c^*) > \frac{2}{3\zeta}(b_j - c^*)$, thereby satisfying Inequality (I.$j+1$). ∎

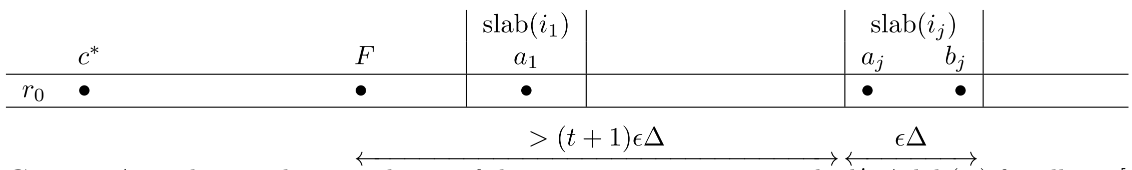
LEMMA 3.2. *Suppose $A(r_0, c^*) = A(r_1, c^*) = 1$ have the same sig_1-type for some indices $r_0 < r_1$ and c^*, and both satisfy Inequality (I.1). Then A contains P_t.*

Proof. Let the landmarks for $A(r_0, c^*)$ and $A(r_1, c^*)$ be $(F, a_1, b_1, \ldots, L)$ and $(F', a'_1, b'_1, \ldots, L')$, respectively. Call $\mathrm{slab}(i)$ the interval of columns $\left[c^* + (1+\epsilon)^i, c^* + (1+\epsilon)^{i+1}\right)$. If the first vector in the common sig_1-signature is $(i_1, \ldots, i_{t-1})$, then $a_j, a'_j \in \mathrm{slab}(i_j)$ for all $j \in [t-1]$. (Since Inequality (I.1) is satisfied, all a_j, a'_j exist.)

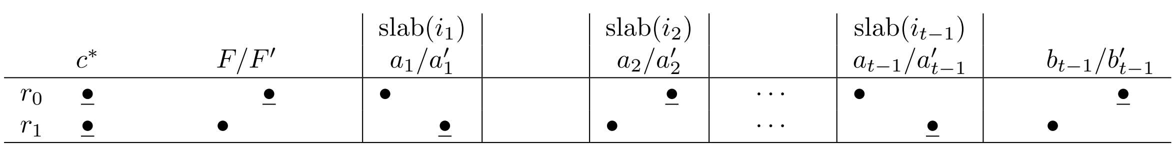
Case 1: Suppose that according to the second part of the common sig_1-signature, there is an index j such that $b_j, b'_j \in \mathrm{slab}(i_j)$. Then $a_j - c^*$ is at least the distance from c^* to $\mathrm{slab}(i_j)$, which is $\Delta \stackrel{\mathrm{def}}{=} (1+\epsilon)^{i_j}$, and $b_j - a_j$ is at most $\epsilon\Delta$, the width of $\mathrm{slab}(i_j)$. Since Inequality (I.1) is satisfied and $a_1 - c^* < (1+\epsilon)^{i_1+1}$, $F - c^* < (1+\frac{1}{3\zeta})^{-1}(1+\epsilon)^{i_1+1}$. Thus,

$$
\begin{aligned}
a_j - F = (a_j - c^*) - (F - c^*) &\geq \Delta - (1+\tfrac{1}{3\zeta})^{-1}(1+\epsilon)^{i_1+1} \\
&\geq \Delta\left(1 - \tfrac{(1+\epsilon)3\zeta}{3\zeta+1}\right) && (i_1 \leq i_j) \\
&> \frac{\Delta}{6(\zeta+1)}. && (\epsilon < 1/(6\zeta))
\end{aligned}
$$

By the definition of a_j, b_j, every interval of width $\epsilon\Delta$ in row r_0 (resp., r_1) between F (resp., F') and $\mathrm{slab}(i_j)$ contains a 1. Thus, rows r_0 and r_1 contain an alternating pattern of length $(6(\zeta+1)\epsilon)^{-1} > t+1$, and together with column c^* this forms an instance of P_t. See the diagram below.



Case 2: According to the second part of the common sig_1-signature, $b_j, b'_j \notin \mathrm{slab}(i_j)$ for all $j \in [t-1]$, hence $i_1 < i_2 < \cdots < i_{t-1}$. Since $a_1 - F > (F - c^*)/(3\zeta)$ and $\epsilon < 1/(3\zeta)$, F is not in $\mathrm{slab}(i_1)$. Then we have an instance of P_t on rows r_0, r_1 and columns $c^* < F < a'_1 < a_2 < \cdots < \{a_{t-1}, a'_{t-1}\} < \{b_{t-1}, b'_{t-1}\}$, where the columns selected from $\{a_{t-1}, a'_{t-1}\}$ and $\{b_{t-1}, b'_{t-1}\}$ depend on the parity of t. The underlined points in the figure below form an instance of P_t if t is even.



Cases 1 and 2 are exhaustive, so we conclude A contains P_t. ∎

LEMMA 3.3. *Suppose there are two rows $r_0 < r_1$ and three columns $c^* < c, c'$ such that*

- *$A(r_0, c^*) = A(r_0, c) = A(r_1, c^*) = A(r_1, c') = 1$ and all have a common sig_2-type.*
- *$A(r_0, c^*)$ and $A(r_1, c^*)$ are each not the last 1 in their row with their respective sig_0-types. They* do not *satisfy Inequality (I.1).*
- *Let the landmarks for $A(r_0, c^*), A(r_1, c^*)$ be $(F, a_1, b_1, \ldots, L)$ and $(F', a'_1, b'_1, \ldots, L')$, respectively. Both $c \in [F, a_1)$ and $c' \in [F', a'_1)$.*

Then A contains P_t.

Proof. Since $c \in [F, a_1)$, the landmarks for $A(r_0, c)$ are $(\tilde{F}, a_1, \ldots, b_{t-1}, L)$, i.e., they only differ from the landmarks for $A(r_0, c^*)$ in $F/\tilde{F}$.

Let the first vector of the common sig_2-signature be $(i_1, \ldots, i_{t-1})$, i.e., if i_j is defined then b_j, b'_j exist and are in $\mathrm{slab}(i_j)$, where the slabs are defined as in the proof of Lemma 3.2. Since $\lfloor \log_{1+\epsilon}(b_1 - c^*) \rfloor = \lfloor \log_{1+\epsilon}(b_1 - c) \rfloor$, this implies $F - c^* \leq c - c^* < \epsilon(1+\epsilon)^{i_1}$ (the width of $\mathrm{slab}(i_1)$) and that F lies strictly before $\mathrm{slab}(i_1)$. We similarly have $F' - c^* < \epsilon(1+\epsilon)^{i_1}$.

Case 1: According to the second part of the common sig_2-signature, for some index $j \in [t-1]$, a_j, a'_j exist and are in $\mathrm{slab}(i_j)$. Then $b_j - a_j < \epsilon(1+\epsilon)^{i_j} \stackrel{\mathrm{def}}{=} \epsilon\Delta$ and by definition of a_j, b_j, every interval of width $\epsilon\Delta$ in row r_0 between F and $\mathrm{slab}(i_j)$ has a 1. The same is true in row r_1 (with F' in place of F), so there is an alternating pattern of length $(\Delta - \epsilon(1+\epsilon)^{i_1})/\epsilon\Delta \geq (1-\epsilon)/\epsilon > t+1$ between rows r_0, r_1, and together with column c^*, this forms an instance of P_t.

From now on we assume we are not in Case 1, so in particular, we have $i_1 < i_2 < \cdots < i_{t-1}$.

Case 2: According to the common sig_2-signature, for some index $j \in [t-1]$, a_j is strictly between $\mathrm{slab}(i_{j-1})$ and $\mathrm{slab}(i_j)$. Then A contains an instance of P_t on rows r_0, r_1 and the $t+2$ columns $c^*, F, b'_1, \cdots, \{b_{j-1}, b'_{j-1}\}, \{a_j, a'_j\}, \{b_j, b'_j\}, \ldots, \{b_{t-1}, b'_{t-1}\}$, where the columns selected from $\{a_j, a'_j\}, \{b_{t-1}, b'_{t-1}\}$, etc. depend on the parities of j and t.

Case 3: According to the common sig_2-signature, for every $j \in [2, t-1]$, $a_j, a'_j \in \mathrm{slab}(i_{j-1})$. By Lemma 3.1, at least one of Inequalities (I.1)–(I.t) are satisfied, but Inequality (I.1) is not satisfied by assumption. Since $a_j - b_{j-1}$ is less than the width of $\mathrm{slab}(i_{j-1})$, we cannot satisfy any of Inequalities (I.2)–(I.$t-1$), hence Inequality (I.t) is satisfied: $L - b_{t-1} > \frac{1}{3\zeta}(b_{t-1} - c^*)$, implying that L lies outside $\mathrm{slab}(i_{t-1})$ since $\epsilon < 1/(3\zeta)$. Thus, A contains an instance of P_t on rows r_0, r_1 and the $t+2$ columns $c^*, F, b'_1, \ldots, \{b_{t-1}, b'_{t-1}\}, \{L, L'\}$. ∎

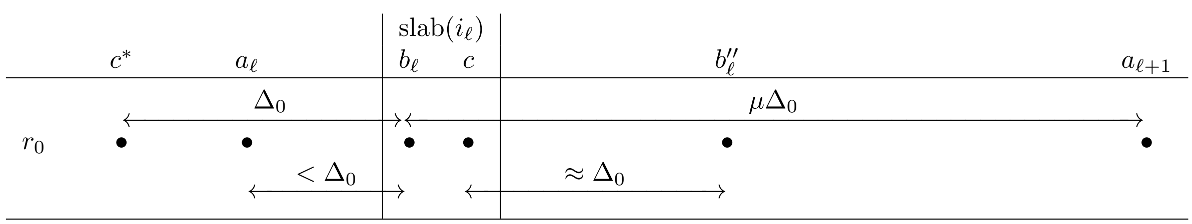
LEMMA 3.4. *Suppose there are two rows $r_0 < r_1$ and three columns $c^* < c, c'$ such that*

- $A(r_0, c^*) = A(r_0, c) = A(r_1, c^*) = A(r_1, c') = 1$ *and all have a common* sig_2*-type.*
- *Let the landmarks for $A(r_0, c^*), A(r_1, c^*)$ be $(F, a_1, b_1, \ldots, L)$ and $(F', a'_1, b'_1, \ldots, L')$, respectively (with the first few a_i, b_i potentially undefined). For some $\ell \in [t-1]$, $c \in [b_\ell, a_{\ell+1})$ and $c' \in [b'_\ell, a'_{\ell+1})$, where $a_t \stackrel{\mathrm{def}}{=} L, a'_t \stackrel{\mathrm{def}}{=} L'$.*

Then A contains P_t.

Proof. Since $c \in [b_\ell, a_{\ell+1})$, the landmarks for $A(r_0, c)$ are $(F'', a''_1, \ldots, b''_\ell, a_{\ell+1}, \ldots, a_{t-1}, b_{t-1}, L)$, i.e., they agree on the suffix $(a_{\ell+1}, \ldots, L)$ with the landmarks of $A(r_0, c^*)$. Let $(i_1, \ldots, i_{t-1})$ be the first component of the common sig_2-signature. Define $\Delta_0 \in [(1+\epsilon)^{i_\ell}, (1+\epsilon)^{i_\ell+1})$ and $\mu > 0$ as:

$$\Delta_0 = b_\ell - c^*,$$
$$\mu\Delta_0 = a_{\ell+1} - b_\ell.$$



The relevant component of the third part of the common sig_2-signature is $\min\left\{\left\lfloor \frac{a_{\ell+1}-b_\ell}{(b_\ell-c^*)/2}\right\rfloor, 3t\right\} = \min\{\lfloor 2\mu \rfloor, 3t\}$. Observe that because of the first part of the common sig_2-signature, $b''_\ell - c \in ((1+\epsilon)^{-1}\Delta_0, (1+\epsilon)\Delta_0)$. Thus,

$$\frac{a_{\ell+1} - b''_\ell}{b''_\ell - c} \leq \frac{\mu\Delta_0}{(1+\epsilon)^{-1}\Delta_0} - 1 = (1+\epsilon)\mu - 1.$$

Note that since $\epsilon < 1/(3t)$, $(1+\epsilon)\mu - 1 < \mu - 1/2$ whenever $\mu < 3t/2$. Since the sig_2-signatures of $A(r_0, c^*), A(r_0, c)$ are identical, it must be that

$$\min\{\lfloor 2\mu \rfloor, 3t\} = \min\left\{\left\lfloor \frac{a_{\ell+1} - b_\ell}{(b_\ell - c^*)/2} \right\rfloor, 3t\right\} = \min\left\{\left\lfloor \frac{a_{\ell+1} - b''_\ell}{(b''_\ell - c)/2} \right\rfloor, 3t\right\} = 3t.$$

Since $b_\ell - a_\ell < \Delta_0$ is maximum among distances of consecutive 1s in the range $[F, a_{\ell+1}]$, there must be a 1 in row r_0 in every interval of width Δ_0 that starts in the range $[a_\ell, a_{\ell+1}] \supset [c^* + (1+\epsilon)^{i_\ell}, c^* + (\mu+1)(1+\epsilon)^{i_\ell}]$. The same is true of row r_1 with respect to some $\Delta_1 \in [(1+\epsilon)^{i_\ell}, (1+\epsilon)^{i_\ell+1})$. Since $\mu \geq 3t/2 \geq t+1$, there are $t+1$ alternations between rows r_0, r_1 in the interval $[c^* + (1+\epsilon)^{i_\ell}, c^* + (\mu+1)(1+\epsilon)^{i_\ell}]$. Together with column c^* this forms an instance of P_t. □

We are now in a position to prove Theorem 3.1, using Lemmas 3.1 to 3.4.

Proof. [Proof of Theorem 3.1] Let A be an $n \times n$ matrix. Recall the 4-step marking process.

Step 1. In each row, mark the last two 1s and the last 1 of each sig_0-type.

Step 2. In each column, mark the last unmarked 1 of each sig_1-type that satisfies Inequality (I.1).

Step 3. In each row, mark the last t unmarked 1s of each sig_2-type.

Step 4. In each column, mark the last t unmarked 1s of each sig_2-type.

We shall prove that either every 1 is marked, or A contains an instance of P_t. Suppose that there exists some $A(r_0, c^*) = 1$ that is still unmarked after Steps 1–4. Because it is unmarked after Step 1, Lemma 3.1 implies that it must satisfy at least one of Inequalities (I.1)–(I.t). If it satisfies Inequality (I.1) then Step 2 must have marked some other $A(r_1, c^*) = 1$, $r_1 \neq r_0$, of the same sig_1-type that also satisfies Inequality (I.1). In this case A contains P_t, by Lemma 3.2. Thus, after Step 2, we may assume that $A(r_0, c^*)$ does not satisfy Inequality (I.1).

Steps 3 and 4 imply that there exists rows $r_0 < r_1 < \cdots < r_t$ and for each $i \in [0, t]$, columns $c^* < c_{i,1} < c_{i,2} < \cdots < c_{i,t}$ such that $A(r_i, c^*) = A(r_i, c_{i,j}) = 1$ all have the common sig_2-type of $A(r_0, c^*)$. Each $A(r_i, c_{i,j})$ is classified by which of the following sets contains $c_{i,j}$, the landmarks $(F^i, a_1^i, b_1^i, \ldots, L^i)$ being defined w.r.t. $A(r_i, c^*)$.

$$[F^i, a_1^i),\ \{a_1^i\},\ [b_1^i, a_2^i),\ \{a_2^i\},\ [b_2^i, a_3^i)\ \ldots,\ \{a_{t-1}^i\},\ [b_{t-1}^i, L^i).$$

There are only $t-1$ singleton sets $\{a_1^i\}, \ldots, \{a_{t-1}^i\}$, so by the pigeonhole principle, for each $i \in [0, t]$ there exists a $j(i) \in [t]$ such that $c_{i,j(i)} \in [F^i, a_1^i) \cup [b_1^i, a_2^i) \cup \cdots \cup [b_{t-2}^i, a_{t-1}^i) \cup [b_{t-1}^i, L^i)$. By the pigeonhole principle again, there must exist two rows $r_i, r_{i'}$ such that $c_{i,j(i)}, c_{i',j(i')}$ have the same classification. If $c_{i,j(i)} \in [F^i, a_1^i)$ and $c_{i',j(i')} \in [F^{i'}, a_1^{i'})$ then Lemma 3.3 implies that A contains P_t. If, for some $\ell \in [t-1]$, $c_{i,j(i)} \in [b_\ell^i, a_{\ell+1}^i)$ and $c_{i',j(i')} \in [b_\ell^{i'}, a_{\ell+1}^{i'})$ (by definition $a_t^i \stackrel{\text{def}}{=} L^i$), then Lemma 3.4 implies A contains P_t.

The cases considered above are exhaustive, hence if A is P_t-free, there can be no unmarked 1s left in A after Steps 1–4.

The number of distinct sig_0-, sig_1-, and sig_2-signatures are $O(n \log_\zeta^t n), O(n \log_{1+\epsilon}^{t-1} n)$, and $O(n \log_{1+\epsilon}^{t-1} n)$, respectively, so the number of 1s marked by Steps 1–4 is

$$O(n \cdot (\log_\zeta^t n + \log_{1+\epsilon}^{t-1} n)) = O\left(n \cdot \left((\log n / \log \zeta)^t + (\zeta \log n)^{t-1}\right)\right). \qquad \text{(Recall } \epsilon = \Theta(1/\zeta).\text{)}$$

We choose $\zeta = (\log n)^{1/t}$ and conclude that $\text{Ex}(P_t, n) = O(n(\log n / \log\log n)^t)$. □

Remark 3.1. *Note that if A is a rectangular $n \times m$ matrix, Steps 1 and 3 mark $O(n(\log_\zeta^t m + \log_{1+\epsilon}^{t-1} m))$ 1s whereas Steps 2 and 4 mark $O(m \log_{1+\epsilon}^{t-1} m)$ 1s. If $n > m$ then we still pick $\zeta = (\log m)^{1/t}$ and get the upper bound $\text{Ex}(P_t, n, m) = O(n(\log m / \log\log m)^t)$, whereas if $n < m/\log m$, we would pick $\zeta = O(1)$ and get the upper bound $\text{Ex}(P_t, m/\log m, m) = O(m \log^{t-1} m)$. This shows another qualitative difference between the behavior of P_1-free and P_t-free, $t \geq 2$, rectangular matrices. Pettie [Pet10, Appendix A] proved that for $n \geq m$, $\text{Ex}(P_1, n, m) = \Theta(n \log_{1+n/m} m)$ and for $m \geq n$, $\text{Ex}(P_1, n, m) = \Theta(m \log_{1+m/n} n)$. I.e., whenever $\max\{n/m, m/n\} = \text{poly}(\log(nm))$, we only lose a $\Theta(\log\log n)$-factor in the density of P_1-free matrices, but can lose an $\Omega(\log n / \text{poly}(\log\log n))$-factor in the density of P_t-free matrices.*

4 Conclusion

The foremost open problem in forbidden 0–1 patterns is *still* to understand the range of possible extremal functions for acyclic patterns. In the past it has been valuable to advance conjectures of *varying strength* (implausibility) [FH92, PT06]. Here we present *weak* and *strong* variants of the central conjecture.

Conjecture 4.1. *Let $\mathcal{P}_{acyclic}$ be the class of acyclic 0–1 patterns.*

Weak Form. *For all $P \in \mathcal{P}_{acyclic}$, $\mathrm{Ex}(P,n) = n^{1+o(1)}$. (Cf. [Tar05, PT06].)*

Strong Form. *For all $P \in \mathcal{P}_{acyclic}$, there exists a constant C_P such that $\mathrm{Ex}(P,n) = O(n2^{C_P\sqrt{\log n}})$.*

One way to begin to tackle the **Weak Form** of Conjecture 4.1 is to prove that there exists an absolute constant $\epsilon > 0$ such that $\mathrm{Ex}(P,n) = O(n^{2-\epsilon})$, for all $P \in \mathcal{P}_{\mathrm{acyclic}}$. Another is to generalize Korándi et al.'s [KTTW19] method to handle patterns like T_1 (which can be decomposed using *both* vertical and horizontal cuts) and ultimately to patterns like the pretzel T_0, for which there is no horizontal/vertical cut that separates the 1s in a single column/row. The **Strong Form** of the conjecture should be considered more plausible in light of [KT23a].

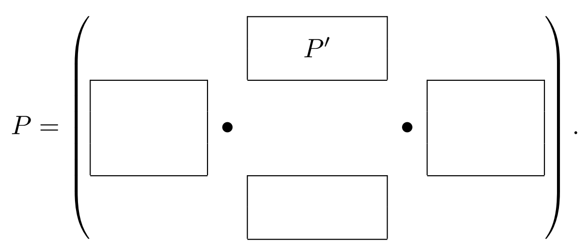
Regardless of the status of Conjecture 4.1, it would be of great interest to characterize the **Polylog** echelon: those P with $\mathrm{Ex}(P,n) = O(n\log^{C_P} n)$.

The *applications* of extremal 0–1 matrix theory in combinatorics [Kla00, MT04], graph theory [GM14, BGK$^+$21, BKTW22], and data structures [Pet10, CGK$^+$15b, CGJ$^+$23, CPY24] tend to use patterns that are rather low in the hierarchy, usually **Linear** or **Quasilinear**, so there is ample motivation to understand the boundary between these two echelons.

Conjecture 4.2. *Let $\mathcal{Q}$ contain Q_3, Q_3', and their horizontal and vertical reflections. If P is light and $\mathcal{Q}$-free then $\mathrm{Ex}(P,n) = O(n)$.*

$$Q_3 = \begin{pmatrix} \bullet & & \bullet & \\ & \bullet & & \bullet \end{pmatrix}, \qquad Q_3' = \begin{pmatrix} \bullet & & & \\ & & \bullet & \\ & \bullet & & \bullet \end{pmatrix}.$$

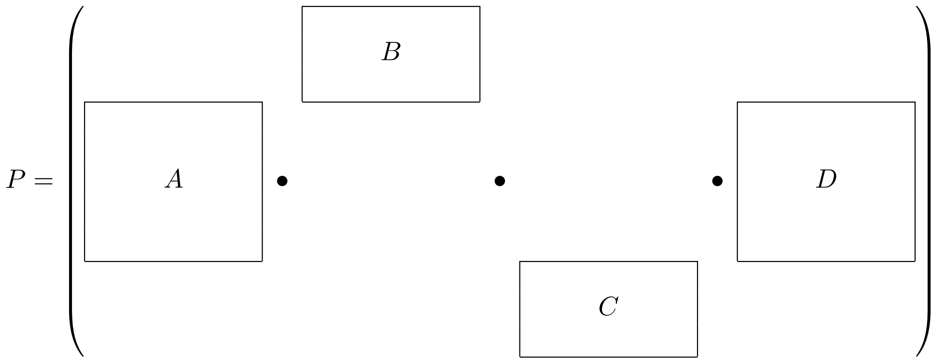
Conjecture 4.2 has not even been confirmed for all weight-5 light patterns, though most have been classified; see [Tar05, Lemma 5.1], Fulek [Ful09, Theorem 4], Pettie [Pet11b, Theorem 2.3], and [MT04, Pet15b, CPY24]. One way to think about classifying light $\mathcal{Q}$-free P is to ignore any consecutive repeated columns in P, then measure how many rows have at least two 1s. If this number is zero then P is a permutation matrix (possibly with repeated columns) and $\mathrm{Ex}(P,n) = O(n)$ [MT04, Gen09]. If P has one such row, then the remaining 1s must be arranged in a permutation P' constrained to the boxed regions in the figure below.



At the other extreme, there are $\mathcal{Q}$-free patterns in which all but one row contain two non-consecutive 1s, but they are all contained in the oscillating patterns (O_t). Only O_2 is known to be linear [FH92]. Fulek [Ful09] proved that the version of O_3 without the repeated column is linear.

$$O_2 = \begin{pmatrix} & \bullet & \bullet & \\ \bullet & & & \bullet \end{pmatrix}, \quad O_3 = \begin{pmatrix} & \bullet & & & \bullet & \\ \bullet & & & & & \bullet \\ & & \bullet & \bullet & & \end{pmatrix}, \quad O_4 = \begin{pmatrix} & & & \bullet & \bullet & & & \\ & \bullet & & & & & \bullet & \\ \bullet & & & & & & & \bullet \\ & & \bullet & & & \bullet & & \end{pmatrix}.$$

There is a class of light $\mathcal{Q}$-free P containing a row with *three* non-consecutive 1s, but patterns in this class are highly constrained in their structure. Each must be of the following form, up to reflection.



where $B, C \neq 0$, A, D are permutations avoiding the weight-3 row, and B, C are $\begin{pmatrix} \bullet & & \bullet \\ & \bullet & \end{pmatrix}$-free and $\begin{pmatrix} & \bullet & \\ \bullet & & \bullet \end{pmatrix}$-free, respectively. In particular, there cannot be a second row (intersecting A, B, C, or D) that has three non-consecutive 1s. There are no light $\mathcal{Q}$-free patterns containing four non-consecutive 1s.

References

[AFK$^+$24] Amir Abboud, Nick Fischer, Zander Kelley, Shachar Lovett, and Raghu Meka. New graph decompositions and combinatorial boolean matrix multiplication algorithms. In *Proceedings 56th ACM Symposium on Theory of Computing (STOC)*, 2024.

[Beh46] Felix A. Behrend. On sets of integers which contain no three terms in arithmetic progression. *Proc. Nat. Acad. Sci.*, 32:331–332, 1946.

[BG91] Daniel Bienstock and Ervin Győri. An extremal problem on sparse 0–1 matrices. *SIAM J. Discr. Math.*, 4(1):17–27, 1991.

[BGK$^+$21] Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width II: small classes. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1977–1996, 2021.

[BKTW22] Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width I: tractable FO model checking. *J. ACM*, 69(1):3:1–3:46, 2022.

[BS74] John A. Bondy and Miklós Simonovits. Cycles of even length in graphs. *J. Combinatorial Theory, Series B*, 16:97–105, 1974.

[BS23] Thomas F. Bloom and Olof Sisask. The Kelley–Meka bounds for sets free of three-term arithmetic progressions. *Essential Number Theory*, 2(1):15–44, 2023.

[CGJ$^+$23] Parinya Chalermsook, Manoj Gupta, Wanchote Jiamjitrak, Nidia Obscura Acosta, Akash Pareek, and Sorrachai Yingchareonthawornchai. Improved pattern-avoidance bounds for greedy BSTs via matrix decomposition. In *Proceedings of the 34th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 509–534, 2023.

[CGK$^+$15a] Parinya Chalermsook, Mayank Goswami, László Kozma, Kurt Mehlhorn, and Thatchaphol Saranurak. Greedy is an almost optimal deque. In *Proceedings 14th Symposium on Algorithms and Data Structures*, volume 9214 of *Lecture Notes in Computer Science*, pages 152–165, 2015.

[CGK$^+$15b] Parinya Chalermsook, Mayank Goswami, László Kozma, Kurt Mehlhorn, and Thatchaphol Saranurak. Pattern-avoiding access in binary search trees. In *Proceedings of the 56th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 410–423, 2015.

[Cib09] Josef Cibulka. On constants in the Füredi-Hajnal and the Stanley-Wilf conjecture. *J. Comb. Theory, Ser. A*, 116(2):290–302, 2009.

[CK17] Josef Cibulka and Jan Kynčl. Better upper bounds on the Füredi-Hajnal limits of permutations. In *Proceedings of the 28th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2280–2293, 2017.

[CPY24] Parinya Chalermsook, Seth Pettie, and Sorrachai Yingchareonthawornchai. Sorting pattern-avoiding permutations via 0-1 matrices forbidding product patterns. In *Proceedings 35th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 133–149, 2024.

[ES46] Paul Erdős and Arthur H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[ES66] Paul Erdős and Miklós Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar.*, 1:51–57, 1966.

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

[FJK+21] Zoltán Füredi, Tao Jiang, Alexandr Kostochka, Dhruv Mubayi, and Jacques Verstraëte. Extremal problems for convex geometric hypergraphs and ordered hypergraphs. *Canadian Journal of Mathematics*, 73(6):1648–1666, 2021.

[FKMV20] Zoltán Füredi, Alexandr V. Kostochka, Dhruv Mubayi, and Jacques Verstraëte. Ordered and convex geometric trees with linear extremal function. *Discret. Comput. Geom.*, 64(2):324–338, 2020.

[Fox13] Jacob Fox. Stanley-Wilf limits are typically exponential, 2013. arXiv:1310.8378.

[FS13] Zoltán Füredi and Miklós Simonovits. The history of degenerate (bipartite) extremal graph problems. In László Lovász, Imre Z. Ruzsa, and Vera T. Sós, editors, *Erdős Centennial*, number 25 in Bolyai Society Mathematical Studies, pages 169–264. Springer, 2013.

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

[Für90] Zoltán Füredi. The maximum number of unit distances in a convex n-gon. *J. Combin. Theory Ser. A*, 55(2):316–320, 1990.

[Gen09] Jesse T. Geneson. Extremal functions of forbidden double permutation matrices. *J. Combin. Theory Ser. A*, 116(7):1235–1244, 2009.

[Gen15] Jesse T. Geneson. *Bounds on extremal functions of forbidden patterns*. PhD thesis, Massachusetts Institute of Technology, Department of Mathematics, 2015.

[Gen19] Jesse T. Geneson. Forbidden formations in multidimensional 0–1 matrices. *Eur. J. Comb.*, 78:147–154, 2019.

[GKM+18] Ervin Györi, Dániel Korándi, Abhishek Methuku, István Tomon, Casey Tompkins, and Máté Vizer. On the Turán number of some ordered even cycles. *Eur. J. Comb.*, 73:81–88, 2018.

[GM14] Sylvain Guillemot and Dániel Marx. Finding small patterns in permutations in linear time. In *Proceedings 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 82–101, 2014.

[GMN+23] Dániel Gerbner, Abhishek Methuku, Dániel T. Nagy, Dömötör Pálvölgyi, Gábor Tardos, and Máté Vizer. Turán problems for edge-ordered graphs. *J. Comb. Theory, Ser. B*, 160:66–113, 2023.

[GT17] Jesse T. Geneson and Peter M. Tian. Extremal functions of forbidden multidimensional matrices. *Discret. Math.*, 340(12):2769–2781, 2017.

[GT20] Jesse T. Geneson and Shen-Fu Tsai. Sharper bounds and structural results for minimally nonlinear 0-1 matrices. *Electron. J. Comb.*, 27(4):4, 2020.

[HS86] Sergiu Hart and Micha Sharir. Nonlinearity of Davenport-Schinzel sequences and of generalized path compression schemes. *Combinatorica*, 6(2):151–177, 1986.

[JJMM24] Barnabás Janzer, Oliver Janzer, Van Magnan, and Abhishek Methuku. Tight general bounds for the extremal numbers of 0–1 matrices. *International Mathematics Research Notices*, to appear.

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

[Kla92] Martin Klazar. A general upper bound in extremal theory of sequences. *Comment. Math. Univ. Carolin.*, 33(4):737–746, 1992.

[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, Berlin, 2000.

[Kla02] Martin Klazar. Generalized Davenport-Schinzel sequences: results, problems, and applications. *Integers*, 2:A11, 2002.

[KM23] Zander Kelley and Raghu Meka. Strong bounds for 3-progressions. In *Proceedings 64th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 933–973, 2023.

[KT23a] Gaurav Kucheriya and Gábor Tardos. A characterization of edge-ordered graphs with almost linear extremal functions. *Combinatorica*, 43:1111–1123, 2023.

[KT23b] Gaurav Kucheriya and Gábor Tardos. On edge-ordered graphs with linear extremal functions. In *Proceedings of the European Conference on Combinatorics, Graph Theory, and Applications (EUROCOMB)*, pages 688–694, 2023.

[KTTW19] Dániel Korándi, Gábor Tardos, István Tomon, and Craig Weidert. On the Turán number of ordered forests. *J. Comb. Theory, Ser. A*, 165:32–43, 2019.

[KV94] Martin Klazar and Pavel Valtr. Generalized Davenport-Schinzel sequences. *Combinatorica*, 14(4):463–476, 1994.

[Mit92] Joseph S. B. Mitchell. L_1 shortest paths among polygonal obstacles in the plane. *Algorithmica*, 8(1):55–88, 1992.

[MT04] Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley-Wilf conjecture. *J. Combin. Theory Ser. A*, 107(1):153–160, 2004.

[MT22] Abhishek Methuku and István Tomon. Bipartite Turán problems for ordered graphs. *Comb.*, 42(6):895–911, 2022.

[Niv10] Gabriel Nivasch. Improved bounds and new techniques for Davenport-Schinzel sequences and their generalizations. *J. ACM*, 57(3), 2010.

[Pet08] Seth Pettie. Splay trees, Davenport-Schinzel sequences, and the deque conjecture. In *Proceedings 19th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1115–1124, 2008.

[Pet10] Seth Pettie. Applications of forbidden 0–1 matrices to search tree- and path compression-based data structures. In *Proceedings 21st ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1457–1467, 2010.

[Pet11a] Seth Pettie. Degrees of nonlinearity in forbidden 0–1 matrix problems. *Discrete Mathematics*, 311:2396–2410, 2011.

[Pet11b] Seth Pettie. Generalized Davenport-Schinzel sequences and their 0–1 matrix counterparts. *J. Comb. Theory Ser. A*, 118(6):1863–1895, 2011.

[Pet11c] Seth Pettie. On the structure and composition of forbidden sequences, with geometric applications. In *Proceedings 27th Annual Symposium on Computational Geometry (SoCG)*, pages 370–379, 2011.

[Pet15a] Seth Pettie. Sharp bounds on Davenport-Schinzel sequences of every order. *J. ACM*, 62(5):36, 2015.

[Pet15b] Seth Pettie. Three generalizations of Davenport-Schinzel sequences. *SIAM J. Discrete Mathematics*, 29(4):2189–2238, 2015.

[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.

[PS13] Sung Gi Park and Qinru Shi. New bounds on extremal numbers in acyclic ordered graphs. https://math.mit.edu/news/summer/SPURprojects/2013ParkShi.pdf, 2013.

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

[PT24] Seth Pettie and Gábor Tardos. On the extremal functions of acyclic forbidden 0-1 matrices. In *Proceedings 35th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1166–1176, 2024.

[SA95] Micha Sharir and Pankaj Agarwal. *Davenport-Schinzel Sequences and their Geometric Applications*. Cambridge University Press, 1995.

[SY17] Asaf Shapira and Raphael Yuster. A tournament approach to pattern avoiding matrices. *Israel Journal of Mathematics*, 217:477–505, 2017.

[Tar05] Gábor Tardos. On 0–1 matrices and small excluded submatrices. *J. Combin. Theory Ser. A*, 111(2):266–288, 2005.

[Tur41] Paul Turán. On an extremal problem in graph theory. *Matematikai és Fisikai Lapok*, 48:436–452, 1941.

[WP18] Julian Wellman and Seth Pettie. Lower bounds on Davenport-Schinzel sequences via rectangular Zarankiewicz matrices. *Discret. Math.*, 341(7):1987–1993, 2018.

[WS88] Ady Wiernik and Micha Sharir. Planar realizations of nonlinear Davenport-Schinzel sequences by segments. *Discrete Comput. Geom.*, 3(1):15–47, 1988.

A Refutation of the Füredi et al. Conjecture for Hypergraphs

We start with defining some simple terms. For the standerd definition of *vertex-ordered hypergraph* and *order-isomorphism*, see [FJK+21]. Define $\mathrm{Ex}^r_<(H,n)$ to be the maximum number of edges in an r-uniform, vertex-ordered hypergraph not containing any subgraphs order-isomorphic to an r-uniform, vertex-ordered H. The *interval chromatic number* of H, denoted $\chi_<(H)$, is the smallest number of intervals that $V(H)$ can be partitioned into (w.r.t. the order on $V(H)$) such that each hyperedge intersects r distinct intervals. As in the Erdős-Stone-Simonovits theorem [ES46, ES66], $\mathrm{Ex}^r_<(H,n) = \Theta(n^r)$ if $\chi_<(H) > r$. A hypergraph H is called a *forest* if there is a peeling order $e_1,\ldots,e_{|E(H)|}$ where for all $i < |E(H)|$, there exists an $h > i$ such that $e_i \cap \bigcup_{j>i} e_j \subset e_h$.

A closely related concept is the generalization of extremal function Ex from 2-dimensional matrices to r-dimensional matrices. If $P \in \{0,1\}^{k_1\times\cdots\times k_r}$, $A \in \{0,1\}^{n_1\times\cdots\times n_r}$, we say $P \prec A$ if there are index sets (I_i), $I_i \subseteq [n_i]$, $|I_i| = k_i$, such that P is entry-wise dominated by the submatrix $A[I_1,\ldots,I_r]$ of A restricted to $I_1,\ldots,I_r$. Define $\mathrm{Ex}^r(P,n) = \max\{\|A\|_1 \mid A \in \{0,1\}^{n^r} \text{ and } P \not\prec A\}$. We can view P as an ordered hypergraph $H(P)$ with interval chromatic number $\chi_<(H(P)) = r$ by ordering the $k_1+\cdots+k_r$ vertices primarily by their axis, then according to their coordinate within the axis. Each $P(i_1,\ldots,i_r) = 1$ corresponds to a hyperedge $\{i_1^{(1)},\ldots,i_r^{(r)}\}$, where $i_j^{(j)}$ is the vertex in axis j and position i_j.

Lemma A.1. (Cf. [PT06, Thm. 2]) *Suppose $P \in \{0,1\}^{k_1\times\cdots\times k_r}$ and $H(P)$ has no isolated vertices. Then $\mathrm{Ex}^r(P,n) \le \mathrm{Ex}^r_<(H(P),rn)$.*

Proof. If $A \in \{0,1\}^{r^n}$ is P-free, then $H(A)$ is an ordered r-uniform hypergraph on rn vertices that is $H(P)$-free, as every embedding of $H(P)$ in $H(A)$ must assign the r vertices of each hyperedge into vertices associated with distinct axes of A. ∎

Recall that Conjecture 1.4 is equivalent to the $r = 2$ case of Conjecture 1.5. Thus, Theorem 1.2 refuting the former gives a counterexample of the latter conjecture that is a 2-uniform ordered hypergraph, i.e., a vertex-ordered graph. We thought that providing counterexamples that are "real" hypergraphs is valuable, so here we give a similar counterexample that is an r-uniform ordered hypergraphs for any $r > 2$, as claimed in the first paragraph of Section 1.6.

Let $S_0^r \in \{0,1\}^{3\times 4\times 1\times\cdots\times 1}$ be the r-dimensional version of the pattern S_0, where the last $r-2$ dimensions have width 1. Then $H(S_0^r)$ is in fact an ordered r-uniform forest hypergraph with $r+5$ vertices. Construct a matrix $A^r \in \{0,1\}^{n^r}$ such that every square submatrix $A \in \{0,1\}^{n\times n}$ obtained by fixing a single coordinate in the last $r-2$ axes is a copy of the lower bound construction A from Theorem 1.2. Then $\mathrm{Ex}^r(S_0^r, n) \geq n^{r-2}\|A\|_1 = n^{r-1}2^{\sqrt{\log n}-O(\log\log n)}$. By Lemma A.1, $\mathrm{Ex}^r_<(H(S_0^r), n) \geq \mathrm{Ex}^r(S_0^r, n/r) = \Omega(n^{r-1}2^{\sqrt{\log n}-O(\log\log n)})$, which refutes Conjecture 1.5 for all $r \geq 2$ as $H(S_0^r)$ is an r-uniform forest of interval chromatic number r.