ARTICLE IN PRESS

European Journal of Combinatorics xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Combinatorics

journal homepage: https://www.sciencedirect.com/journal/european-journal-of-combinatorics

Posets are easily testable

Panna Tímea Fekete *, Gábor Kun

Institute of Mathematics, Eötvös Loránd University, POB 120, H-1518 Budapest, Hungary
HUN-REN Alfréd Rényi Institute of Mathematics, POB 127, H-1364 Budapest, Hungary

ARTICLE INFO

Article history:
Available online xxxx

ABSTRACT

Alon and Shapira proved that every monotone class (closed under taking subgraphs) of undirected graphs is strongly testable, that is, under the promise that a given graph is either in the class or ε-far from it, there is a test using a constant number of samples (depending on ε only) that rejects every graph not in the class with probability at least one half, and always accepts a graph in the class. However, their bound on the number of samples is quite large since they heavily rely on Szemerédi's regularity lemma. We study the case of posets and show that every monotone class of posets is easily testable, that is, a polynomial (of ε^{-1}) number of samples is sufficient. We achieve this via proving a polynomial removal lemma for posets.

We give a simple classification: for every monotone class of posets, there is an h such that the class is indistinguishable (every large enough poset in one class is ε-close to a poset in the other class) from the class of C_h-free posets, where C_h denotes the chain with h elements. This allows us to test every monotone class of posets using $O(\varepsilon^{-1})$ samples. The test has a two-sided error, but it is almost complete: the probability of refuting a poset in the class is polynomially small in the size of the poset.

The analogous results hold for comparability graphs, too.

1. Introduction

The relationship between local and global properties of structures is a central theme in combinatorics and computer science. Since the work of Rubinstein and Sudan [24], testing properties by

* Corresponding author at: Institute of Mathematics, Eötvös Loránd University, POB 120, H-1518 Budapest, Hungary.
E-mail address: fekete.panna.timea@renyi.hu (P.T. Fekete).

https://doi.org/10.1016/j.ejc.2024.104044

sampling a small number of elements is an emerging research area. A classical result of this kind is the triangle removal lemma by Ruzsa and Szemerédi [25], usually stated in the form that if a graph G admits at most $\delta|V(G)|^3$ triangles then it can be made triangle-free by the removal of at most $\varepsilon|V(G)|^2$ edges, where δ depends only on ε. This can be applied to obtain a combinatorial proof of Roth's theorem [23] on 3-term arithmetic progressions, while the hypergraph removal lemma has been used to prove Szemerédi's theorem. Removal lemmas were proved for abelian groups by Green [16], for linear systems of equations by Král, Serra and Vena [20], for local affine-invariant properties by Bhattacharyya, Fischer, Hatami, Hatami and Lovett [9] and for permutations by Klimošová and Král [19], and by Fox and Wei [11], as well.

A *property* of digraphs is a set of finite digraphs closed under isomorphism. A digraph G is ε*-far* from having a property Φ if any digraph G' on the vertex set $V(G)$ that differs by at most $\varepsilon|V(G)|^2$ edges from G does not have the property Φ either. A property Φ is *strongly testable* if for every $\varepsilon > 0$ there exists an $f(\varepsilon)$ such that if the digraph G is ε-far from having the property Φ then the induced directed subgraph on $f(\varepsilon)$ vertices chosen uniformly at random does not have the property Φ with probability at least one half, and it always has the property if G does. Alon and Shapira [5] proved that every *monotone property* of undirected graphs (that is, closed under the removal of edges and vertices) is strongly testable, see Lovász and Szegedy for an analytic approach [21], while Rödl and Schacht generalized this to hypergraphs [22], see also Austin and Tao [7]. Similar results have been obtained for hereditary classes of graphs and other structures, e.g., tournaments and matrices, see Gishboliner for the most recent summary [12]. We focus on monotone properties and omit the overview of other research directions.

Unfortunately, the dependence on ε can be quite bad already in the case of undirected graphs: the known upper bounds in the Alon-Shapira theorem are wowzer functions due to the iterated involvement of Szemerédi's regularity lemma. Following Alon and Fox [3], we call a property *easily testable* if $f(\varepsilon)$ can be bounded by a polynomial of $\frac{1}{\varepsilon}$, else the property is *hard*. They showed that both testing perfect graphs and testing comparability graphs are hard [2]. Easily testable properties are quite rare, even triangle-free graphs are hard: Behrend's construction [8] of sets of integers without 3-term arithmetic progression leads to a lower bound of magnitude $\varepsilon^{c\log\left(\frac{1}{\varepsilon}\right)}$. Alon proved that H-freeness is easily testable in the case of undirected graphs if and only if H is bipartite. For forbidden induced subgraphs, Alon and Shapira gave a characterization [6], where there are very few easy cases. Testability is usually hard for hypergraphs studied by Gishboliner and Shapira [13] and ordered graphs investigated by Gishboliner and Tomon [14]. An interesting class of properties that are easy to test are semialgebraic hypergraphs, see Fox, Pach and Suk [10]. Surprisingly, 3-colorability and, in general, "partition problems" turned out to be easily testable, see Goldreich, Goldwasser and Ron [15]. Even a conjecture to draw the borderline between easy and hard properties seems beyond reach.

The goal of this paper is to study testability of finite posets as special digraphs. By a poset, we mean a set equipped with a partial order $\prec$ that is anti-reflexive and transitive. Alon, Ben-Eliezer and Fischer [1] proved that hereditary (closed under induced subgraphs) classes of ordered graphs are strongly testable. This implies the removal lemma for posets and that monotone classes of posets are strongly testable in the following way. We consider a linear extension $<$ of the ordering $\prec$ of the poset P. To every poset with a linear ordering, we can associate the graph on its base set, where distinct elements $x < y$ are adjacent if $x \prec y$ in the poset. A graph with a linear ordering is associated with a poset if and only if it has no induced subgraph with two edges on three vertices, where the smallest and largest vertices are not adjacent. An alternative to the application of this general result is to follow the proof of Alon and Shapira [5] using the poset version of Szemerédi's regularity lemma proved by Hladký, Máthé, Patel and Pikhurko [17].

We show that monotone classes of posets (closed under taking subposets) are easily testable. This is equivalent to the following removal lemma with polynomial bounds.

Throughout this paper, we work with finite posets. The *height* of a poset P is the length of its longest chain, while the *width* is the size of the largest antichain, denoted by $h(P)$ and $w(P)$, respectively. The chain with h elements is denoted by C_h. Given two posets P, Q, a mapping $f : Q \to P$ is a homomorphism if it is order-preserving, i.e., $f(x) \prec f(y)$ for every $x \prec y$. The probability that a uniform random mapping from Q to P is a homomorphism is denoted by $t(Q, P)$, which we often refer to as the homomorphism density. A poset P is called Q*-free* if it does not contain Q as a (not necessarily induced) subposet.

Theorem 1.1 (*Polynomial Removal Lemma for Posets*)**.** *Consider an* $\varepsilon > 0$ *and a finite poset* Q *of height at least two. For every finite poset* P*, if* $t(Q,P) < \left(\frac{\varepsilon}{2}\right)^{h(Q)w(Q)^2}$ *then there exists a* Q*-free (moreover,* $C_{h(Q)}$*-free) subposet of* P *obtained by the removal of at most* $\varepsilon|P|^2$ *edges.*

We show that Q-free posets are easily testable.

Algorithm 1 Basic test for Q-free posets

Input: the poset P
$P' \leftarrow$ induced subposet on $|Q|$ elements chosen uniformly at random
if Q is a subposet of P' **then** Reject P
else Accept P
end if

This test always accepts a Q-free poset, and rejects a poset P with probability at least $t(Q,P)$. If $t(Q,P) < \left(\frac{\varepsilon}{2}\right)^{h(Q)w(Q)^2}$, then by Theorem 1.1 P is not ε-far from being Q-free. If $t(Q,P) \geq \left(\frac{\varepsilon}{2}\right)^{h(Q)w(Q)^2}$, then it is sufficient to iterate this test $\frac{1}{t(Q,P)} \leq \left(\frac{2}{\varepsilon}\right)^{h(Q)w(Q)^2}$ times independently (i.e., taking $f(\varepsilon) = \left(\frac{2}{\varepsilon}\right)^{h(Q)w(Q)^2}|Q|$ in the definition of easy testability) to reject a poset ε-far from being Q-free with probability at least $1-(1-t(Q,P))^{\frac{1}{t(Q,P)}} > \frac{1}{2}$. The inequality holds since $0 < t(Q,P) \leq 1$, the function $t \mapsto 1-(1-t)^{\frac{1}{t}}$ is monotone increasing on $(0,1]$ and $\lim_{t\to 0} 1-(1-t)^{\frac{1}{t}} = 1-\frac{1}{e}$.

We will consider the family of (possibly infinitely many) finite posets not in the class. To state our precise result, we define the height and width of a set of posets $\mathcal{P}$ as

$$h(\mathcal{P}) = \min_{P\in\mathcal{P}} h(P) \qquad w(\mathcal{P}) = \min_{\substack{P\in\mathcal{P}:\\ h(P)=h(\mathcal{P})}} w(P).$$

Corollary 1.2 (*Easy Testability for Monotone Classes of Posets*)**.** *Consider a family of finite posets* $\mathcal{P}$ *with* $h(\mathcal{P}) \geq 2$*. Let* $Q \in \mathcal{P}$ *with height* $h(Q) = h(\mathcal{P})$ *and width* $w(Q) = w(\mathcal{P})$*. For every* $\varepsilon > 0$ *and finite poset* P*, if* $t(Q,P) < \left(\frac{\varepsilon}{2}\right)^{h(\mathcal{P})w(\mathcal{P})^2}$ *then there exists a* $\mathcal{P}$*-free (moreover,* $C_{h(\mathcal{P})}$*-free) subposet of* P *obtained by the removal of at most* $\varepsilon|P|^2$ *edges.*

Observe that by Theorem 1.1 there exists a $C_{h(Q)}$-free subposet of P obtained by the removal of at most $\varepsilon|P|^2$ edges. Since every poset in $\mathcal{P}$ contains $C_{h(\mathcal{P})}$, this subposet is also $\mathcal{P}$-free, hence Corollary 1.2 holds.

Chains will play an important role in more efficient tests for monotone classes of posets: we give a simple classification of these classes from the testing point of view. Two properties Φ_1 and Φ_2 of posets are *indistinguishable* if for every $\varepsilon > 0$ and $i = 1,2$ there exists N such that for every poset P on at least N elements with property Φ_i there exists a poset P' on the same set with property Φ_{3-i} obtained by changing at most $\varepsilon|P|^2$ edges of P. Since we are interested in monotone properties, we only need to allow deleting edges and not adding them.

Theorem 1.3 (*Indistinguishability*)**.** *Consider a family of finite posets* $\mathcal{P}$*, set* $h = h(\mathcal{P}) \geq 2$ *and* $w = w(\mathcal{P})$*. The class of* $\mathcal{P}$*-free posets and the class of* C_h*-free posets are indistinguishable. Namely, every* C_h*-free poset is* $\mathcal{P}$*-free, and if a poset* P *is* $\mathcal{P}$*-free then it has a* C_h*-free subposet obtained by the removal of at most* $2\left(\frac{h^2w^2}{|P|}\right)^{\frac{1}{hw^2}}|P|^2$ *edges.*

In other words, for every $\mathcal{P}$-free poset P on at least $N = h^2w^2(\varepsilon/2)^{-hw^2}$ elements there exists a C_h-free (not necessarily induced) subposet P' obtained by the removal of at most $\varepsilon|P|^2$ edges.

Theorem 1.3 motivates a better understanding of the removal lemma for chains and the testing of C_h-free posets. First, we study the basic test with one-sided error. We can also use this test for C_h-free posets to test $\mathcal{P}$-free posets, where $h = h(\mathcal{P})$. This test is not complete, but the probability of rejecting a $\mathcal{P}$-free poset turns out to be negligible, $2\left(\frac{h^2w^2}{|P|}\right)^{\frac{1}{hw^2}} \cdot \binom{h}{2}$, where $w = w(\mathcal{P})$, since every copy of C_h should contain one of the edges removed in Theorem 1.3. If we iterate the test $\left(\frac{2}{\varepsilon}\right)^h$ times independently, then the probability of accepting a poset ε-far from being $\mathcal{P}$-free is at most one half by Theorem 1.1. On the other hand, the probability of rejecting a poset that is $\mathcal{P}$-free is at most $2\left(\frac{h^2w^2}{|P|}\right)^{\frac{1}{hw^2}}\binom{h}{2}\left(\frac{2}{\varepsilon}\right)^h$, and this is negligible if ε, h, w are fixed and $|P|$ is large enough.

We can get a more efficient test by sampling larger subposets instead of iterating the basic test with a constant number of samples.

Algorithm 2 Subposet test for C_h-free posets with s samples

Input: the poset P
$P' \leftarrow$ induced subposet of s elements chosen uniformly at random
if C_h is a subposet of P' **then** Reject P
else Accept P
end if

It turns out that sampling $s = \left\lceil \frac{4\log(h)+4}{2\varepsilon} \right\rceil$ elements is enough to reject posets ε-far from being C_h-free with probability at least one half, while we always accept C_h-free posets.

By Theorem 1.3 this test can also be used for testing $\mathcal{P}$-free posets, where $h(\mathcal{P}) = h$: it rejects posets ε-far from $\mathcal{P}$-free with probability at least one half at the price of allowing the error of rejecting a $\mathcal{P}$-free poset with negligible probability.

Theorem 1.4 (*The Subposet Test*). *Let $h \geq 2$ be an integer, $\varepsilon > 0, c > 0$ and P a finite poset. If P is ε-far from being C_h-free then a random subset of $\left\lceil \frac{4\log(h)+4c+1}{2\varepsilon} \right\rceil$ elements chosen independently and uniformly at random contains a copy of C_h with probability at least $1 - e^{-c}$.*

Observe that being ε-far from every C_h-free poset guarantees that ε is small, so the number of samples will be large enough.

Remark 1.5. Every poset P is $\frac{1}{2h-2}$-close to be C_h-free.

Proof. Every poset can be extended to a linear ordering. Partition the poset P into $(h-1)$ intervals of equal size and remove the edges inside the intervals: this gives a C_h-free poset $\frac{1}{2h-2}$-close to P. □

For any fixed h, our bound gives the right order of magnitude (in ε) on the necessary number of samples for one-sided testing of C_h-free posets, see Proposition 2.4.

The comparability graph G associated with a poset P has vertex set $V(G) = P$ and edge set $E(G) = \{(x, y) : x \prec y \text{ or } y \prec x\}$. Alon and Fox proved that it is hard to test if a given graph is a comparability graph [3]. However, under the promise that the input graph is a comparability graph, we can test monotone classes, even though we do not know the underlying poset. All of our results apply to testing monotone classes of comparability graphs, see Section 4.

In a subsequent work, we prove that the exact degree is $(h-1)$ in the polynomial removal lemma for chains (and many other structures). Proposition 2.2 shows that this is sharp. The proof is too technical for this paper to detail here.

In Section 2, we prove the polynomial removal lemma for chains and Theorem 1.4. Section 3 contains the proofs of Theorems 1.1 and 1.3. Section 4 discusses our results on comparability graphs.

2. Testing chains

First, we prove a removal lemma for chains.

Lemma 2.1 (*Removal Lemma for Chains*). *For every $\varepsilon > 0$, positive integer $h \geq 2$ and every finite poset P, if $t(C_h, P) < \left(\frac{\varepsilon}{2}\right)^h$ then there exists a C_h-free subposet of P obtained by the removal of at most $\varepsilon|P|^2$ edges of P.*

Polynomial removal lemmas for directed paths have already been obtained by Alon and Shapira [4], but their bound is $O\left(\varepsilon^{h^2}\right)$. We could use their result to get a removal lemma for chains with a worse polynomial bound. However, we improve their bound to degree h. This is almost the exact degree, as the following example shows.

Proposition 2.2. *Consider the integer $h \geq 2$ and $\varepsilon > 0$ such that ε^{-1} is an integer. Let P be a poset that consists of ε^{-1} chains of equal size at least h and divisible by $(h-1)$.*

(1) *Every subposet obtained by the removal of less than $\frac{1}{2}\left(\frac{\varepsilon}{h-1}|P|^2 - |P|\right)$ edges from P contains C_h as a subposet, hence P is at least $\left(\frac{\varepsilon}{2h-2} - \frac{1}{2|P|}\right)$-far from being C_h-free.*
(2) *The inequality $t(C_h, P) < \frac{\varepsilon^{h-1}}{h!}$ holds.*

Proof. (1) The comparability graph of P is the union of ε^{-1} complete graphs $K_{\varepsilon|P|}$. If P' is a C_h-free subposet of P, then the corresponding comparability graph is K_h-free. By Turán's theorem we have to remove at least $(h-1)\binom{\frac{\varepsilon|P|}{h-1}}{2}$ edges from $K_{\varepsilon|P|}$ in order to obtain a K_h-free graph. Now (1) follows, since $\varepsilon^{-1}(h-1)\binom{\frac{\varepsilon|P|}{h-1}}{2} = \frac{\varepsilon}{2h-2}|P|^2 - \frac{1}{2}|P|$.

(2) The probability that all of the h elements are mapped to the same chain is ε^{h-1}. Note that any homomorphism $C_h \to P$ maps C_h onto an h element chain in P, since P is anti-reflexive. The conditional probability that such a bijection preserves the order of the elements is $\frac{1}{h!}$. □

We consider a linear extension $<$ of the ordering $\prec$ of the poset P. We may assume that the set of elements of P is $[|P|] = \{1, 2 \ldots, |P|\}$, and the linear ordering $<$ is the ordering of the integers.

The algorithm defines a rank function r on the set of elements, such that if $r(y) = k+1$ for some element y, then it has 'many' predecessors $x \prec y$ with $r(x) = k$. Hence, it has 'many' chains C_{k+1} ending at y.

Algorithm 3 Rank function r

Input: $\gamma > 0$, poset P on $[|P|]$, where if $x \prec y$ then $x < y$
  for $y = 1, \ldots, |P|$ **do**
    if $\exists k : \left|\{x : x \prec y, r(x) = k\}\right| \geq \gamma|P|$ **then**
      $r(y) \leftarrow 1 + \max\left\{k : \left|\{x : x \prec y, r(x) = k\}\right| \geq \gamma|P|\right\}$
    else
      $r(y) \leftarrow 1$
    end if
  end for
Output: Rank function $r : P \to \mathbb{Z}_+$

Algorithm 4 will remove the edges to get a C_h-free poset (see Fig. 1 for an example).

Analysis of Algorithm 4:

Claim 2.3. *The following holds.*

(1) *The output P' is a poset.*
(2) *The output poset P' is C_h-free.*

Algorithm 4 Edge removal using the rank function r

Input: $\gamma > 0$, $h \in \mathbb{Z}_+$, poset P on $[|P|]$, where if $x \prec y$ then $x < y$
  ALGORITHM 3(γ, P)
  for $x \prec y$ **do**
    if $r(x) = r(y)$ **then**
      $E(P) \leftarrow E(P) \setminus \{(x, y)\}$
    else if $r(y) \geq h$ **then**
      $E(P) \leftarrow E(P) \setminus \{(x, y)\}$
    end if
  end for
  $P' \leftarrow P$
Output: P' on vertex set $[|P|]$, edge set $E(P') \subseteq E(P)$

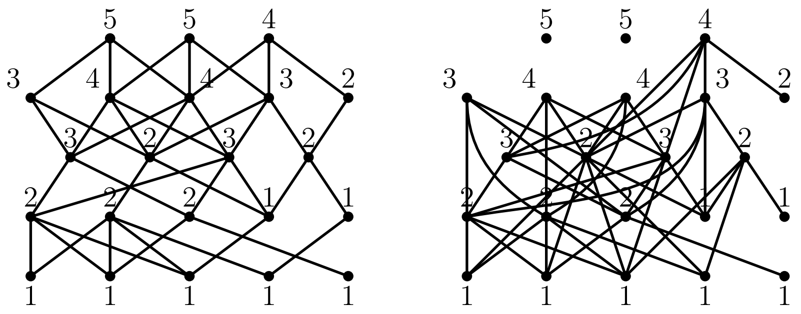

Fig. 1. Example for Algorithm 3 on the left and Algorithm 4 on the right with $h = 5, \gamma = \frac{1}{11}$. The Hasse diagram of the poset P is on the left, the ranks are written on the elements, and the Hasse diagram of P' is on the right.

(3) *The number of edges $x \prec y$ removed such that $r(x) = r(y)$ is at most $\gamma|P|^2$.*
(4) *If the number of elements with rank $r(y) \geq h$ is at most $\gamma|P|$, then the number of edges removed by Algorithm 4 in order to get a C_h-free poset is at most $2\gamma|P|^2$.*

Proof. (1) If x, y, z are distinct elements in P with $(x, y) \in E(P')$ and $(y, z) \in E(P')$, then $(x, z) \in E(P)$, and $r(z) < h$, $r(x) < r(y) < r(z)$. Hence $(x, z) \in E(P')$.

(2) Let x, y be distinct elements in P with $(x, y) \in E(P')$. Note that $r(x) \leq r(y)$ by the transitivity in posets, hence r is non-decreasing on every chain in P. Every edge with $r(x) = r(y)$ has been removed. Thus, r is strictly increasing on every chain in P'. The poset P' is C_h-free since the edges ending at those elements, where r is at least h, have been removed.

(3) For every y, the number of $x \prec y$ with $r(x) = r(y)$ can be at most $\gamma|P|$, else $r(y)$ would be greater than $r(x)$. So, the number of such removed edges is at most $\gamma|P|^2$.

(4) This is a straightforward consequence of the algorithm and (3). □

Proof of Lemma 2.1. We run Algorithm 4 with h, P and $\gamma = \frac{\varepsilon}{2}$.

Claim. *If $t(C_h, P) < \gamma^h$, then the number of elements with rank $r(y) \geq h$ is strictly less than $\gamma|P|$. In particular, there is no element with rank $(h + 1)$.*

Proof. Observe that there are at least $(\gamma|P|)^{r(x)-1}$ chains on $r(x)$ elements ending at x for every x such that r is strictly increasing on these chains.

There is no element where r takes value $(h + 1)$ since such an element would be the end of at least $\gamma|P|^h$ chains on at least $(h + 1)$ elements, but we do not have so many different chains of length h. By the same reason, the number of elements, where r takes value h, is strictly less than $\gamma|P|$. ■

(4) of Claim 2.3 proves the lemma. □

Now we use the rank function defined by Algorithm 3 to optimize the number of samples to test C_h-free posets.

Proof of Theorem 1.4. We consider again a linear extension $<$ of the ordering $\prec$ of the poset P. We might assume that the set of elements of P is $[|P|] = \{1, 2 \dots, |P|\}$, and the linear ordering $<$ is the ordering of the integers. We define $r : P \mapsto \mathbb{Z}_+$ using Algorithm 3 with $\gamma = \frac{\varepsilon}{2}$.

Let X be a subset of $\left\lceil \frac{4\log(h)+4c+1}{2\varepsilon} \right\rceil = \left\lceil \frac{\log(h)+c}{\gamma} + \frac{1}{4\gamma} \right\rceil$ elements chosen uniformly at random from P. We prove that with probability at least $(1-e^{-c})$ there is a chain with elements $x_h \prec \cdots \prec x_2 \prec x_1$ such that $r(x_k) = h - k + 1$ for all $k \in [h]$. We will find these elements one by one, starting with x_1.

We show that there are at least $\gamma|P|$ elements $x \in P$ with $r(x) = h$. Suppose for a contradiction that there are less. Then running Algorithm 4 gives a C_h-free poset and by (4) of Claim 2.3 we removed at most $\varepsilon|P|^2$ edges, contradicting that P was ε-far from being C_h-free.

Thus, the probability that we do not choose any element with $r(x) = h$ into the set X is at most $(1-\gamma)^{\gamma^{-1}(\log(h)+c+1/4)} < \frac{e^{-c}}{h}$. Denote by x_1 the smallest element (in the linear extension) such that $r(x_1) = h$, if there is such an element.

Claim. *Consider $x_1, x_2, \dots, x_k$ for $k < h$ such that for every $\ell \in \{2, 3, \dots, k\}$ the element x_ℓ is the smallest (in the linear extension) such that $r(x_\ell) = h - \ell + 1$ and $x_\ell \prec x_{\ell-1}$. Then the conditional distribution on the choice of $x_1, \dots, x_k$ of the other elements of X is uniform on the set*

$$S_k := \{x \in P \setminus \{x_1, x_2, \dots, x_k\} : \forall \ell \in [k] \text{ if } x < x_\ell \text{ then } \{r(x) \neq h - \ell + 1\} \vee \{x \not\prec x_{\ell-1}\}\}.$$

Proof. Note that $X \subseteq S_k \cup \{x_1, x_2, \dots, x_k\}$: else for the smallest x (in the linear extension) such that $x \notin S_k \cup \{x_1, x_2, \dots, x_k\}$ there would be an ℓ such that $x < x_\ell$, $r(x) = h - \ell + 1$ and $x \prec x_{\ell-1}$. Hence, we should have chosen x instead of x_ℓ.

On the other hand, the set X could be $S' \cup \{x_1, x_2, \dots, x_k\}$ for any subset $S' \subseteq S_k$ of size $\left\lceil \frac{\log(h)+c}{\gamma} + \frac{1}{4\gamma} \right\rceil - k$. Since the conditional distribution of X is uniform on these sets, the claim follows. ■

Now we show that a suitable x_{k+1} exists with probability at least $1 - \frac{e^{-c}}{h}$.

There are at least $\gamma|P|$ elements $x \in P$ (in particular, $x \in P \setminus \{x_1, x_2, \dots, x_k\}$ by the partial ordering) such that $x \prec x_k$ and $r(x) = h - k$ by the definition of the rank function. Let us denote these good candidates for x_{k+1} by R_{k+1}.

Since $\varepsilon < \frac{1}{2h-2}$ and $\gamma < \frac{1}{4h-4}$, there are at least $\frac{\log(h)+c}{\gamma}$ elements in $X \setminus \{x_1, x_2, \dots, x_k\}$. The probability that none of them is in R_{k+1} is at most $(1-\gamma)^{\gamma^{-1}(\log(h)+c)} < \frac{e^{-c}}{h}$. Let x_{k+1} be the smallest element (in the linear extension) such that $x_{k+1} \in R_{k+1} \cap X$ if there is such an element.

The union bound yields the theorem. □

The following proposition shows that Theorem 1.4 gives the right order of magnitude on the number of samples required for one-sided testing.

We denote by $K_{w_1, w_2, \dots, w_k}$ the complete h-partite poset: the set of elements consists of pairwise disjoint antichains A_i of size w_i for $1 \leq i \leq k$; and $x \prec y$ for $x \in A_i$ and $y \in A_j$ if and only if $i < j$. Let $K_{h\times w}$ be the shorthand notation for $K_{w,w,\dots,w}$ with hw elements. In particular, $K_{h\times 1}$ is the chain C_h.

Proposition 2.4. *Given $\varepsilon > 0$ and the positive integers $h \geq 2$, $w \geq 1$ such that εw is also an integer, consider the poset $P = K_{\varepsilon w, w, w, \dots, w}$ with $(\varepsilon + h - 1)w$ elements.*

(1) *Every subposet obtained by the removal of less than εw^2 edges from P contains C_h as a subposet, hence P is at least $\frac{\varepsilon}{(\varepsilon+h-1)^2}$-far from being C_h-free.*

(2) *For any $0 < c < \varepsilon w$ the probability that a random subset with at most $\frac{c}{2\varepsilon}$ elements does not contain C_h as a subposet is at least e^{-c}.*

Note that the bound in (1) is sharp: if we remove all of the εw^2 edges between the first two antichains, we obtain a C_h-free poset.

Proof. (1) Every edge with an endvertex in the first antichain (of size εw) is contained by exactly w^{h-2} chains of height h, since we can choose the other elements of the chain from the other antichains arbitrarily. On the other hand, an edge not adjacent to the first antichain is contained by εw^{h-2} chains. Since P contains εw^h chains of height h, we need at least εw^2 edges to cover these.

(2) Every subposet isomorphic to C_h has an element in the first antichain. The probability that a subposet on $k \leq \frac{c}{2\varepsilon} < \frac{w}{2}$ elements contains no element of this antichain is

$$\prod_{i=1}^{k} \frac{(h-1)w - i + 1}{(h-1+\varepsilon)w - i + 1} > \left(\frac{1}{1+2\varepsilon}\right)^{\frac{c}{2\varepsilon}} > e^{-c}. \quad \square$$

This gives the right order of magnitude of the number of samples required for the one-sided testing of C_h-free posets for every fixed h: Theorem 1.4 shows that using $\left\lceil \frac{4\log(h)+4c+1}{2\varepsilon} \right\rceil$ samples the error probability is at most e^{-c}, while Proposition 2.4 gives an example where the error is at least e^{-c} when sampling at most $\frac{c}{2\varepsilon}$ elements.

3. Testing monotone classes of posets

The following lemma provides a lower bound on the density of the complete h-partite poset $K_{h\times w}$ in terms of the density of the chain of length h. The proof is inspired by the counting argument of Kővári, Sós and Turán [18] in the proof of the upper bound to the symmetric case of the Zarankiewicz problem (that is, using modern notation, the upper bound on $ex(n, K_{r,r})$). We use again the notation $[n] = \{1, 2, \ldots, n\}$.

Lemma 3.1. *For every poset P and positive integers h, w the inequality*

$$t(K_{h\times w}, P) \geq t^{w^2}(C_h, P)$$

holds.

Proof. The following two claims imply the lemma.

Claim.

$$t(K_{w,1,w,1,\ldots}, P) \geq t^{w}(C_h, P)$$

Proof. Note that $K_{w,1,w,1,\ldots}$ is the union of w edge-disjoint chains of length h intersecting only on the elements of the even layers (where it has only one element), and a mapping of $K_{w,1,w,1,\ldots}$ is a homomorphism if and only if its restriction to every chain is a homomorphism. Consider a mapping of the even layers of $K_{w,1,w,1,\ldots}$. The events that the random mapping gives a homomorphism for chains are conditionally independent for disjoint chains (conditioning on the mapping of the even layers). Hence, the conditional probability that mapping w elements for every odd layer gives a homomorphism of $K_{w,1,w,1,\ldots}$ is the wth power of the probability that mapping only one element for every odd layer gives a homomorphism of the chain C_h. We use Jensen's inequality to obtain the required result. Now, we describe this argument more formally.

Let $(x_{i,j})_{i\in[h], j\in[w] \text{ for } i \text{ odd}}$ where $x_{i,j} \in P$ and $(x_{i,1})_{i\in[h] \text{ for } i \text{ even}}$ be chosen uniformly and independently at random in P.

$$t(K_{w,1,w,1,\ldots}, P) = \mathbb{P}_{\substack{(x_{i,1})_{i\in[h]} \text{ for } i \text{ even} \\ (x_{i,j})_{i\in[h],j\in[w]} \text{ for } i \text{ odd}}} \left(\forall k \in [h-1], \ell \in [w] \quad \begin{matrix} \text{if } k \text{ odd then } x_{k,\ell} \prec x_{k+1,1} \\ \text{if } k \text{ even then } x_{k,1} \prec x_{k+1,\ell} \end{matrix}\right)$$

$$= \mathbb{E}_{\substack{(x_{i,1})_{i\in[h]} \\ i \text{ even}}} \left[\mathbb{P}_{\substack{(x_{i,j})_{i\in[h],j\in[w]} \\ i \text{ odd}}} \left(\forall k \in [h-1], \ell \in [w] \quad \begin{matrix} \text{if } k \text{ odd then } x_{k,\ell} \prec x_{k+1,1} \\ \text{if } k \text{ even then } x_{k,1} \prec x_{k+1,\ell} \end{matrix} \,\middle|\, (x_{i,1})_{i\in[h]}, i \text{ even}\right)\right]$$

Here we split $K_{w,1,w,1,\ldots}$ into w edge-disjoint copies of C_h. Since the events corresponding to elements in the same odd layer are independent, we obtain that this equals

$$\begin{aligned}
&\mathbb{E}_{\substack{(x_{i,1})_{i\in[h]}\\ i \text{ even}}}\left[\mathbb{P}_{\substack{(x_{i,1})_{i\in[h]}\\ i \text{ odd}}}\left(\forall k\in[h-1]\quad x_{k,1}\prec x_{k+1,1}\,\middle|\,(x_{i,1})_{i\in[h]},\, i \text{ even}\right)\right]^w\\
&\geq\left[\mathbb{E}_{\substack{(x_{i,1})_{i\in[h]}\\ i \text{ even}}}\mathbb{P}_{\substack{(x_{i,1})_{i\in[h]}\\ i \text{ odd}}}\left(\forall k\in[h-1]\quad x_{k,1}\prec x_{k+1,1}\,\middle|\,(x_{i,1})_{i\in[h]},\, i \text{ even}\right)\right]^w\\
&=\left[\mathbb{P}_{(x_{i,1})_{i\in[h]}}\left(\forall k\in[h-1]\quad x_{k,1}\prec x_{k+1,1}\right)\right]^w=t^w(C_h,P),
\end{aligned}$$

where we have applied Jensen's inequality. ■

Claim.

$$t(K_{h\times w},P)\geq t^w(K_{w,1,w,1,\ldots},P)$$

Proof. The proof is very similar to the previous one. Now we use the observation that $K_{h\times w}$ is the union of w edge-disjoint copies of $K_{w,1,w,1,\ldots}$ intersecting only on the odd layers (where $K_{w,1,w,1,\ldots}$ has w elements), and a mapping of $K_{h\times w}$ is a homomorphism if and only if its restriction to every such copy of $K_{w,1,w,1,\ldots}$ is a homomorphism. Consider a mapping of the odd layers of $K_{h\times w}$. The events that the random mapping gives a homomorphism for copies of $K_{w,1,w,1,\ldots}$ are conditionally independent for disjoint copies of $K_{w,1,w,1,\ldots}$ (conditioning on the mapping of the odd layers). Hence, the conditional probability that mapping w elements for every even layer gives a homomorphism of $K_{h\times w}$ is the wth power of the probability that mapping only one element for every even layer gives a homomorphism of $K_{w,1,w,1,\ldots}$. We use Jensen's inequality again to obtain the required result.

Let $(x_{i,j})_{i\in[h],j\in[w]}$ be chosen uniformly and independently at random in P.

$$\begin{aligned}
t(K_{h\times w},P)&=\mathbb{P}_{(x_{i,j})_{i\in[h],j\in[w]}}\left(\forall k\in[h-1],\,\ell,m\in[w]\quad x_{k,\ell}\prec x_{k+1,m}\right)\\
&=\mathbb{E}_{\substack{(x_{i,j})_{i\in[h],j\in[w]}\\ i\text{ odd}}}\left[\mathbb{P}_{\substack{(x_{i,j})_{i\in[h],j\in[w]}\\ i\text{ even}}}\left(\forall k\in[h-1],\,\ell,m\in[w]\quad x_{k,\ell}\prec x_{k+1,m}\,\middle|\,(x_{i,j})_{i\in[h],j\in[w]},\, i\text{ odd}\right)\right].
\end{aligned}$$

Here we split $K_{h\times w}$ into w edge-disjoint copies of $K_{w,1,w,1,\ldots}$. Since the events corresponding to elements in the same even layer are independent, we obtain that this equals

$$\begin{aligned}
&\mathbb{E}_{\substack{(x_{i,j})_{i\in[h],j\in[w]}\\ i\text{ odd}}}\left[\mathbb{P}_{\substack{(x_{i,1})_{i\in[h]}\\ i\text{ even}}}\left(\forall k\in[h-1],\,\ell\in[w]\quad \begin{matrix}\text{if } k \text{ odd then } x_{k,\ell}\prec x_{k+1,1}\\ \text{if } k \text{ even then } x_{k,1}\prec x_{k+1,\ell}\end{matrix}\,\middle|\,(x_{i,j})_{i\in[h],j\in[w]},\, i\text{ odd}\right)\right]^w\\
&\geq\left[\mathbb{E}_{\substack{(x_{i,j})_{i\in[h],j\in[w]}\\ i\text{ odd}}}\mathbb{P}_{\substack{(x_{i,1})_{i\in[h]}\\ i\text{ even}}}\left(\forall k\in[h-1],\,\ell\in[w]\quad \begin{matrix}\text{if } k \text{ odd then } x_{k,\ell}\prec x_{k+1,1}\\ \text{if } k \text{ even then } x_{k,1}\prec x_{k+1,\ell}\end{matrix}\,\middle|\,(x_{i,j})_{i\in[h],j\in[w]},\, i\text{ odd}\right)\right]^w\\
&=\left[\mathbb{P}_{\substack{(x_{i,1})_{i\in[h]}\ \text{for } i\text{ even}\\ (x_{i,j})_{i\in[h],j\in[w]}\ \text{for } i\text{ odd}}}\left(\forall k\in[h-1],\,\ell\in[w]\quad \begin{matrix}\text{if } k \text{ odd then } x_{k,\ell}\prec x_{k+1,1}\\ \text{if } k \text{ even then } x_{k,1}\prec x_{k+1,\ell}\end{matrix}\right)\right]^w=t^w(K_{w,1,w,1,\ldots},P),
\end{aligned}$$

where we have applied Jensen's inequality. ■

The lemma follows. □

Proof of Theorem 1.1. Assume that $t(Q,P)<\left(\frac{\varepsilon}{2}\right)^{hw^2}$. The poset Q is a subposet of $K_{h\times w}$, so Lemma 3.1 gives $t^{w^2}(C_h,P)\leq t(K_{h\times w},P)\leq t(Q,P)$. These yield $t(C_h,P)<\left(\frac{\varepsilon}{2}\right)^h$, so by Lemma 2.1 there is a C_h-free subposet P' of P obtained by deleting at most $\varepsilon|P|^2$ edges. □

Proof of Theorem 1.3. If a poset is C_h-free, then it is $\mathcal{P}$-free.

In order to prove the other direction, consider a poset $Q\in\mathcal{P}$ with minimal height $h=h(\mathcal{P})$ and (amongst these) minimal width $w=w(\mathcal{P})$. If a poset P is Q-free, then there is no injective homomorphism from Q to P. The probability that a $Q \to P$ mapping is not injective is at most $\binom{|Q|}{2}|P|^{-1}$ since a pair of elements in Q should be mapped to the same element in P. Thus, $t(Q,P) \leq |P|^{-1}|Q|^2$. Since $|Q| \leq hw$, Theorem 1.1 shows that one can get a C_h-free subposet of P by the removal of $2(h^2w^2)^{\frac{1}{hw^2}}|P|^{-\frac{1}{hw^2}}|P|^2$ edges. □

4. Comparability graphs

We will obtain the same theorems for monotone classes of comparability graphs as for posets: the difference will only be in the hidden constants. These allow the same tests as for posets. For a fixed finite graph F, the basic test samples $|V(F)|$ vertices and accepts a graph if these do not span an isomorphic copy of F. The following removal lemma shows how many iterations we need to reject comparability graphs ε-far from being F-free with probability at least one half while always accepting F-free comparability graphs. Similarly to posets, given two finite graphs F, G, the probability that a uniform random mapping from F to G is a homomorphism (i.e., edge-preserving) is denoted by $t(F,G)$.

Theorem 4.1 (*Polynomial Removal Lemma for Comparability Graphs*)**.** *Consider an $\varepsilon > 0$ and a finite graph F that is not an independent set. For every finite comparability graph G, if $t(F,G) \leq \left(\frac{\varepsilon}{2}\right)^{\chi(F)\alpha(F)^2}$ then there exists an F-free (moreover, $K_{\chi(F)}$-free) spanning subgraph of G that is a comparability graph, obtained by deleting at most $\varepsilon|V(G)|^2$ edges.*

Proof. The graph F is a subgraph of the multipartite Turán graph T with $\chi(F)$ classes each of size $\alpha(F)$, hence $t(F,G) \geq t(T,G)$. There exists a poset P with comparability graph G. The height of the poset P is exactly the chromatic number of G, and the width of the poset P equals the independence number of G.

Note that $t(T,G) \geq t(K_{\chi(F)\times\alpha(F)}, P)$, since we may assume that T is the comparability graph of $K_{\chi(F)\times\alpha(F)}$, hence every homomorphism of $K_{\chi(F)\times\alpha(F)}$ to P is a comparability-preserving map from T to G, i.e., a graph homomorphism. By Theorem 1.1 there exists a $C_{\chi(F)}$-free subposet of P obtained by deleting at most $\varepsilon|P|^2$ edges, and its comparability graph satisfies the conditions of the theorem. □

Note that we did not need to know the underlying poset P to prove the existence of the desired subgraph of G.

Given a set of (possibly infinitely many) finite graphs $\mathcal{F}$, we define the chromatic number $\chi(\mathcal{F})$ and the independence number $\alpha(\mathcal{F})$ as follows.

$$\chi(\mathcal{F}) = \min_{F\in\mathcal{F}} \chi(F) \qquad \alpha(\mathcal{F}) = \min_{\substack{F\in\mathcal{F}:\\ \chi(F)=\chi(\mathcal{F})}} \alpha(F).$$

Corollary 4.2 (*Easy Testability for Monotone Classes of Comparability Graphs*)**.** *Consider a family of finite graphs $\mathcal{F}$ and a graph $F \in \mathcal{F}$ with chromatic number $\chi(\mathcal{F}) \geq 2$ and independence number $\alpha(\mathcal{F})$. For every $\varepsilon > 0$ and finite comparability graph G, if $t(F,G) \leq \left(\frac{\varepsilon}{2}\right)^{\chi(\mathcal{F})\alpha(\mathcal{F})^2}$ then there exists an $\mathcal{F}$-free (moreover, $K_{\chi(\mathcal{F})}$-free) spanning subgraph of G, that is a comparability graph, obtained by deleting at most $\varepsilon|V(G)|^2$ edges.*

We also give a classification of monotone classes of comparability graphs as we did for posets. Two properties Φ_1 and Φ_2 of graphs are *indistinguishable* if for every $\varepsilon > 0$ and $i = 1, 2$ there exists N such that for every graph G on at least N vertices with property Φ_i there exists a graph G' on the same vertex set with property Φ_{3-i}, obtained by changing at most $\varepsilon|V(G)|^2$ edges of G. Since we are interested in monotone properties, we only need to allow deleting edges.

Theorem 4.3 (*Indistinguishability*)**.** *Consider a family of finite graphs $\mathcal{F}$. Set $\chi = \chi(\mathcal{F}) \geq 2, \alpha = \alpha(\mathcal{F})$. Comparability graphs with chromatic number at most $(\chi - 1)$ are indistinguishable from $\mathcal{F}$-free comparability graphs. Namely, every comparability graph with chromatic number at most $(\chi - 1)$ is*

$\mathcal{F}$-free, and every $\mathcal{F}$-free comparability graph admits a spanning subgraph with chromatic number at most $(\chi - 1)$, that is a comparability graph, obtained by the removal of at most $2\left(\frac{\chi^2\alpha^2}{|V(G)|}\right)^{\frac{1}{\chi\alpha^2}}|V(G)|^2$ edges.

Proof. Clearly, every comparability graph with chromatic number at most $(\chi - 1)$ is $\mathcal{F}$-free. On the other hand, given an $\mathcal{F}$-free comparability graph G, consider a poset P whose comparability graph is G. Theorem 1.3 implies that there is a C_χ-free subposet P' obtained by the removal of at most $2\left(\frac{\chi^2\alpha^2}{|V(G)|}\right)^{\frac{1}{\chi\alpha^2}}|V(G)|^2$ edges. The comparability graph G' of P' is the desired spanning subgraph of G: it is K_χ-free, since P' is C_χ-free. Hence, $\chi(G') \leq \chi - 1$ by the dual of the Dilworth theorem. □

The analog of Algorithm 2 is the test sampling a random set of vertices and accepting the graph if the subgraph spanned by them is K_χ-free. We need the same number of samples as in the case of posets. The following theorem is a straightforward consequence of Theorem 1.4.

Theorem 4.4 (*On the Subgraph Test*)**.** *Let $\chi \geq 2$ be an integer, $\varepsilon > 0, c > 0$ and G a finite comparability graph. If G is ε-far from being K_χ-free then a random subset of $\left\lceil \frac{4\log(\chi)+4c+1}{2\varepsilon} \right\rceil$ vertices chosen independently and uniformly at random contains a copy of K_χ with probability at least $1 - e^{-c}$.*

The comparability graph of the poset in Proposition 2.4 shows that for any fixed h, this bound has the right order of magnitude in ε. As in the case of posets, we can also use the test for $K_{\chi(\mathcal{F})}$-free subgraphs to test a monotone class of comparability graphs $\mathcal{F}$: the probability that we reject an $\mathcal{F}$-free comparability graph is negligible.

Acknowledgments

Panna Tímea Fekete's Project No. 1016492. has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the KDP-2020 funding scheme and supported by the ERC Synergy Grant No. 810115 – DYNASNET. Gábor Kun was supported by the Hungarian Academy of Sciences Momentum Grant no. 2022-58 and the ERC Advanced Grant ERMiD.

References

[1] Noga Alon, Omri Ben-Eliezer, Eldar Fischer, Testing hereditary properties of ordered graphs and matrices, in: 2017 IEEE 58th Annual Symposium on Foundations of Computer Science, FOCS, IEEE, 2017.
[2] Noga Alon, Jacob Fox, Testing perfection is hard, 2011, arXiv preprint arXiv:1110.2828.
[3] Noga Alon, Jacob Fox, Easily testable graph properties, Combin. Probab. Comput. 24 (4) (2015) 646–657.
[4] Noga Alon, Asaf Shapira, Testing subgraphs in directed graphs, in: Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, 2003, pp. 700–709.
[5] Noga Alon, Asaf Shapira, Every monotone graph property is testable, in: Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing, 2005, pp. 128–137.
[6] Noga Alon, Asaf Shapira, A characterization of easily testable induced subgraphs, Combin. Probab. Comput. 15 (6) (2006) 791–805.
[7] Tim Austin, Terence Tao, Testability and repair of hereditary hypergraph properties, Random Structures Algorithms 36 (4) (2010) 373–463.
[8] Felix A. Behrend, On sets of integers which contain no three terms in arithmetical progression, Proc. Natl. Acad. Sci. 32 (12) (1946) 331–332.
[9] Arnab Bhattacharyya, Eldar Fischer, Hamed Hatami, Pooya Hatami, Shachar Lovett, Every locally characterized affine-invariant property is testable, in: Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing, 2013, pp. 429–436.
[10] Jacob Fox, János Pach, Andrew Suk, A polynomial regularity lemma for semialgebraic hypergraphs and its applications in geometry and property testing, SIAM J. Comput. 45 (6) (2016) 2199–2223.
[11] Jacob Fox, Fan Wei, Fast property testing and metrics for permutations, Combin. Probab. Comput. 27 (4) (2018) 539–579.
[12] Lior Gishboliner, Efficient Removal Lemmas and Related Problems, Tel Aviv University, 2020.
[13] Lior Gishboliner, Asaf Shapira, Hypergraph removal with polynomial bounds, 2022, arXiv preprint arXiv:2202.07567.

[14] Lior Gishboliner, István Tomon, Polynomial removal lemmas for ordered graphs, 2021, arXiv preprint arXiv:2110.03577.
[15] Oded Goldreich, Shari Goldwasser, Dana Ron, Property testing and its connection to learning and approximation, J. ACM 45 (4) (1998) 653–750.
[16] Ben Green, A szemerédi-type regularity lemma in abelian groups, with applications, Geom. Funct. Anal. GAFA 15 (2) (2005) 340–376.
[17] Jan Hladkỳ, András Máthé, Viresh Patel, Oleg Pikhurko, Poset limits can be totally ordered, Trans. Amer. Math. Soc. 367 (6) (2015) 4319–4337.
[18] P. Kővári, V.T. Sós, P. Turán, On a problem of zarankiewicz, in: Colloquium Mathematicum, Vol. 3, Polska Akademia Nauk., 1954, pp. 50–57.
[19] Tereza Klimošová, Daniel Král', Hereditary properties of permutations are strongly testable, in: Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, 2014, pp. 1164–1173, Society for Industrial and Applied Mathematics.
[20] Daniel Král', Oriol Serra, Lluís Vena, A removal lemma for systems of linear equations over finite fields, Israel J. Math. 187 (2012) 193–207.
[21] László Lovász, Balázs Szegedy, Testing properties of graphs and functions, 2008, arXiv preprint arXiv:0803.1248.
[22] Vojtěch Rödl, Mathias Schacht, Generalizations of the removal lemma, Combinatorica 29 (4) (2009) 467–501.
[23] Klaus F. Roth, On certain sets of integers, J. Lond. Math. Soc. 28 (1) (1953) 104–109.
[24] Ronitt Rubinfeld, Madhu Sudan, Robust characterizations of polynomials with applications to program testing, SIAM J. Comput. 25 (2) (1996) 252–271.
[25] Imre Ruzsa, Endre Szemerédi, Triple systems with no six points carrying three triangles. Combinatorics (keszthely, 1976), Colloquia Math. Soc. J. Bolyai 18 (939–945) (1978) 2.