



# Hereditary properties of partitions, ordered graphs and ordered hypergraphs

József Balogh<sup>a</sup>, Béla Bollobás<sup>b,c</sup>, Robert Morris<sup>b</sup>

<sup>a</sup> *Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States*

<sup>b</sup> *Department of Mathematical Sciences, The University of Memphis, Memphis, TN 38152, United States*

<sup>c</sup> *Trinity College, Cambridge CB2 1TQ, United Kingdom*

Available online 30 August 2006

---

## Abstract

In this paper we use the Klazar–Marcus–Tardos method (see [A. Marcus, G. Tardos, Excluded permutation matrices and the Stanley–Wilf conjecture, *J. Combin. Theory Ser. A* 107 (2004) 153–160]) to prove that, if a hereditary property of partitions  $\mathcal{P}$  has super-exponential speed, then, for every  $k$ -permutation  $\pi$ ,  $\mathcal{P}$  contains the partition of  $[2k]$  with parts  $\{i, \pi(i) + k\} : i \in [k]\}$ . We also prove a similar jump, from exponential to factorial, in the possible speeds of monotone properties of ordered graphs, and of hereditary properties of ordered graphs not containing large complete, or complete bipartite ordered graphs.

Our results generalize the Stanley–Wilf conjecture on the number of  $n$ -permutations avoiding a fixed permutation, which was recently proved by the combined results of Klazar [M. Klazar, The Füredi–Hajnal conjecture implies the Stanley–Wilf conjecture, in: D. Krob, A.A. Mikhalev, A.V. Mikhalev (Eds.), *Formal Power Series and Algebraic Combinatorics*, Springer, Berlin, 2000, pp. 250–255] and Marcus and Tardos [A. Marcus, G. Tardos, Excluded permutation matrices and the Stanley–Wilf conjecture, *J. Combin. Theory Ser. A* 107 (2004) 153–160]. Our main results follow from a generalization to ordered hypergraphs of the theorem of Marcus and Tardos.

© 2006 Elsevier Ltd. All rights reserved.

---

## 1. Introduction

In this paper we shall prove that a jump from exponential to factorial speed occurs for properties of combinatorial structures of various types. We request the reader's patience while we make the various definitions necessary to state our results.

---

*E-mail addresses:* [jobal@math.uiuc.edu](mailto:jobal@math.uiuc.edu) (J. Balogh), [bollobas@msci.memphis.edu](mailto:bollobas@msci.memphis.edu) (B. Bollobás), [rdmorrsl@memphis.edu](mailto:rdmorrsl@memphis.edu) (R. Morris).

An *ordered hypergraph*  $\mathcal{H} = (V, E, <)$  is a hypergraph—a set of vertices  $V$  and edges  $E \subset \{A : A \subset V, |A| \geq 2\}$ —together with a linear order  $<$  on its vertices. Note that we do not allow edges to be repeated, and that we do not allow edges to consist of a single vertex. An ordered hypergraph  $\mathcal{K} = (U, F, <)$  is an *induced sub-hypergraph* of  $\mathcal{H}$  if  $U \subset V$  (with the induced ordering), and  $F = \{e \cap U : e \in E, |e| \geq 2\}$ .  $\mathcal{K}$  is a *sub-hypergraph* of  $\mathcal{H}$  if  $U \subset V$  (again with the induced ordering), and  $F \subset \{e \cap U : e \in E, |e| \geq 2\}$ . Finally,  $\mathcal{K}$  is *contained* in  $\mathcal{H}$  if there exists a sub-hypergraph  $\mathcal{L} = (U, D, <)$  of  $\mathcal{H}$ , with  $|D| = |F| = t$ , say, and  $f_i \subset d_i$  for each  $i \in [t]$  (where  $D = \{d_1, \dots, d_t\}$  and  $F = \{f_1, \dots, f_t\}$ ).

A collection of ordered hypergraphs is called a *property* if it is closed under order-preserving isomorphisms of the vertex set. A property of ordered hypergraphs  $\mathcal{P}$  is called *hereditary* if it is closed under taking induced sub-hypergraphs; it is called *monotone* if it is closed under taking sub-hypergraphs; and it is called *strongly monotone* if it is closed under containment. Observe that any strongly monotone property is monotone, and any monotone property is hereditary.

An *ordered graph* is a graph together with a linear order  $<$  on its vertices; equivalently, it is an ordered hypergraph in which each edge has size exactly 2. The definitions of hereditary and monotone properties are therefore inherited (note that in this case the definitions of monotone and strongly monotone coincide).

A *partition* of the set  $[n] = \{1, \dots, n\}$  is an (unordered) collection of disjoint, non-empty sets  $\{A_1, \dots, A_t\}$  such that  $A_1 \cup \dots \cup A_t = [n]$ . It is easy to see that a partition may be thought of as an ordered graph in which each component is a clique, or as an ordered hypergraph in which the edges are pairwise disjoint. Thus we obtain the definition of a hereditary property of partitions. Since we have come some distance from the original definition, we remark that if  $P = \{A_1, \dots, A_t\}$  is a partition of  $[n]$ , and  $S$  is a subset of  $[n]$  with elements  $s_1 < \dots < s_k$ , then the *sub-partition of  $P$  induced by  $S$*  is the following partition of  $[k]$ . First let  $\{B_1, \dots, B_t\}$  satisfy  $i \in B_j$  if and only if  $s_i \in A_j$ ; then delete the empty classes. A property of partitions is hereditary if it is closed under taking sub-partitions.

Now, given a property  $\mathcal{P}$  of ordered hypergraphs, we write  $\mathcal{P}_n$  for the collection of distinct (i.e. non-isomorphic) ordered hypergraphs on  $n$  vertices in  $\mathcal{P}$ , and call the function  $n \mapsto |\mathcal{P}_n|$  the *speed* (or unlabelled speed) of  $\mathcal{P}$ . An analogous definition can be made for other combinatorial structures (e.g. graphs, posets, permutations).

We are interested in the (surprising) phenomenon that, for many such structures, only very ‘few’ speeds are possible. More precisely, there often exists a family  $\mathcal{F}$  of functions  $f : \mathbb{N} \rightarrow \mathbb{N}$  and another function  $F : \mathbb{N} \rightarrow \mathbb{N}$  with  $F(n)$  *much* larger than  $f(n)$  for every  $f \in \mathcal{F}$ , such that if, for each  $f \in \mathcal{F}$ , the speed is infinitely often larger than  $f(n)$ , then it is also larger than  $F(n)$  for every  $n \in \mathbb{N}$ . Put concisely: the speed *jumps* from  $\mathcal{F}$  to  $F$ .

The study of the speeds of monotone properties of labelled graphs was introduced over 40 years ago by Erdős [11], and continued by Erdős et al. [13], Erdős et al. [12], Kolaitis et al. [21], Kleitman and Winston [20], Hundack et al. [15] and, more recently, Balogh et al. [4]. A new direction was initiated by Scheinerman and Zito [26], who were the first to study hereditary properties of graphs with speeds below  $n^n$ . A little later, considerably stronger results were proved by Balogh et al. [5,6]. In the range  $|\mathcal{P}_n| = 2^{cn^2}$ , the main results were proved by Alekseev [1], Bollobás and Thomason [8,9], and Prömel and Steger [23–25]. For a review of the early results, see Bollobás [7]. Hereditary properties of other combinatorial structures have not yet been studied in such great detail, but it is likely that many more beautiful theorems await discovery.

In this paper we shall prove that a jump of this type, from exponential to factorial speed, occurs for strongly monotone properties of ordered hypergraphs. As a result of this, we shall be

able to prove similar jumps for hereditary properties of partitions, monotone properties of ordered graphs, and hereditary properties of ordered graphs not containing arbitrarily large complete, or complete bipartite ordered graphs. As we shall see, each of these theorems is a generalization of the Stanley–Wilf conjecture ([Theorem A](#)), proved recently by the combined results of Klazar [16] and Marcus and Tardos [22].

Before we begin, we should remark that our main theorem has been proved independently (and at around the same time) by Klazar and Marcus [19]. Although we were unaware of their work until after ours was completed, we should note also that many of the ideas in this paper were inspired by the earlier work of Klazar [16] and of Marcus and Tardos [22]. The reader may also wish to refer to some other papers of Klazar [17,18] which we later discovered contain some of the ideas (though none of the results) below.

For each  $n \in \mathbb{N}$ , let  $\Pi_n$  denote the collection of all permutations of  $[n]$ , and let  $\Pi = \bigcup_n \Pi_n$ . Also, if  $\mathcal{P}$  is a property of ordered hypergraphs, and  $k, \ell \in \mathbb{N} \cup \{\infty\}$ , let

$$\mathcal{P}^{(k,\ell)} = \{G \in \mathcal{P} : \Delta(G) \leq k \text{ and } |e| \leq \ell \text{ for every } e \in E(G)\}$$

denote the sub-property consisting of the ordered hypergraphs in which each vertex is contained in at most  $k$  edges, and each edge has size of at most  $\ell$ . Note that if  $\mathcal{P}$  is hereditary, or monotone, or strongly monotone, then so is  $\mathcal{P}^{(k,\ell)}$ .

Finally, if  $\pi \in \Pi_k$ , let  $H(\pi)$  denote the ordered hypergraph on vertex set  $[2k]$  and with edge set  $\{\{i, \pi(i) + k\} : i \in [k]\}$ . We shall also write  $H(\pi)$  for the ordered graph with the same vertex and edge sets, and for the partition  $\{\{i, \pi(i) + k\} : i \in [k]\}$  of  $[2k]$ . It will always be clear which of these  $H(\pi)$  is.

We are now ready to state the main result of this paper, which was conjectured by Klazar in [18] and has been proved independently by Klazar and Marcus [19].

**Theorem 1.** *Let  $\mathcal{P}$  be a strongly monotone property of ordered hypergraphs. If, for every constant  $c > 0$ , there exists an  $N = N(c) \in \mathbb{N}$  such that  $|\mathcal{P}_N| > c^N$ , then  $\mathcal{P}$  contains the ordered hypergraph  $H(\pi)$  for every  $\pi \in \Pi$ , and hence*

$$|\mathcal{P}_n| \geq |\mathcal{P}_n^{(1,2)}| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k! = n^{n/2+o(n)}$$

for every  $n \in \mathbb{N}$ . This lower bound is the best possible, and there is a unique strongly monotone property of ordered hypergraphs with this speed.

We remark that Scheinerman and Zito [26] proved that a similar jump, from exponential to factorial speed, exists for hereditary properties of labelled graphs. For more involved results, see [5].

Our proof of [Theorem 1](#) is based on the ideas of Klazar [16,18], and of Marcus and Tardos [22]. [Theorem A](#), below, was proved in two stages: first Klazar [16] showed that the theorem was a consequence of a conjecture of Füredi and Hajnal [14]; then Marcus and Tardos [22] proved that conjecture. In [Theorem 2](#), we shall prove a generalization of the theorem of Marcus and Tardos ([Theorem B](#)). We shall then deduce [Theorem 1](#) using the method of Klazar (see [16,18]).

We shall state [Theorem 2](#) in terms of  $(0, 1)$ -matrices, though it can equally be thought of as a theorem about ordered hypergraphs. To simplify the statement, we need a little notation.

Let  $k \in \mathbb{N}$ . If  $A = (a_{ij})$  and  $B = (b_{ij})$  are  $k \times k$  matrices, we shall write  $(A, B)$  for the  $k \times 2k$  matrix obtained by putting  $A$  in front of  $B$ . Thus  $(A, B)_{ij}$  is  $a_{ij}$  if  $j \leq k$  and  $b_{i(j-k)}$  if  $j \geq k+1$ .

Call two matrices  $C$  and  $D$  equivalent (and write  $C \sim D$ ) if  $D$  is obtained from  $C$  by permuting its rows. Let  $\mathcal{M}(k)$  denote the set of equivalence classes (with respect to  $\sim$ ) in the family of all matrices of the form  $(K, L)$ , where  $K$  and  $L$  are  $k \times k$  permutation matrices. Note that every such matrix  $(K, L)$  is equivalent to a unique matrix  $(I, M)$ , where  $I = (\delta_{ij})$  is the  $k \times k$  identity matrix, so  $|\mathcal{M}(k)| = k!$ .

Finally, if  $P$  and  $Q$  are  $(0, 1)$ -matrices, then we say that  $Q$  is a *sub-matrix* of  $P$  if  $Q$  is obtained from  $P$  by deleting rows and columns. We say that  $P$  *contains*  $Q = (q_{ij})$  if there exists a sub-matrix  $R = (r_{ij})$  of  $P$ , the same size as  $Q$ , with  $r_{ij} = 1$  whenever  $q_{ij} = 1$ . If we associate an ordered hypergraph  $\mathcal{H}$  with a  $(0, 1)$ -matrix whose rows are the indicator functions of the edges of  $\mathcal{H}$ , and consider two matrices to be the same if they are equivalent, then this concept coincides with hypergraph containment defined above. If  $P$  does not contain  $Q$ , then we say that  $P$  *avoids*  $Q$ .

The following theorem is a reformulation of Klazar's conjecture C4 in [18]. It has also been proved independently by Klazar and Marcus [19].

**Theorem 2.** *Let  $k \in \mathbb{N}$ . There exists a constant,  $c_k$ , such that, if  $m, n \in \mathbb{N}$  and  $A$  is an  $m \times n$   $(0, 1)$ -matrix satisfying that*

- (i) *at least  $c_k n$  of the entries of  $A$  are 1, and*
- (ii) *each of the rows of  $A$  are different,*

*then  $A$  contains some member of each class of  $\mathcal{M}(k)$ .*

Notice that Theorem 2 still holds if condition (ii) is replaced by the condition

- (ii') *each of the rows of  $A$  has at least  $2k$  of its entries 1,*

since, if  $A$  satisfies (ii'), and any row occurs at least  $k$  times in  $A$ , then  $A$  contains every  $k \times 2k$   $(0, 1)$ -matrix.

**Remark 1.** Füredi and Hajnal [14] proved that the extremal number of 1s possible in an  $n \times n(0, 1)$ -matrix avoiding the matrix

$$S_1 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

is (up to a constant)  $n\alpha(n)$ , where  $\alpha(n) \rightarrow \infty$  extremely slowly. A simple corollary of Theorem 2 (with  $k = 2$ ) is that the extremal number of 1s if we avoid both  $S_1$  and

$$S_2 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

is  $O(n)$ . For many more results along these lines, see Tardos [27].

We shall now note two important (and immediate) consequences of Theorem 1. The first of them was conjectured by Klazar in [17], and the second was originally proved (although not stated!) by Klazar in [16] as a consequence of the Füredi–Hajnal Conjecture.

**Theorem 3.** *Let  $\mathcal{P}$  be a hereditary property of partitions. If, for every constant  $c > 0$ , there exists an  $N = N(c) \in \mathbb{N}$  such that  $|\mathcal{P}_N| > c^N$ , then  $\mathcal{P}$  contains the partition  $H(\pi)$  for every  $\pi \in \Pi$ , and hence*

$$|\mathcal{P}_n| \geq |\mathcal{P}_n^{(1,2)}| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k! = n^{n/2+o(n)}$$

for every  $n \in \mathbb{N}$ . This lower bound is best possible, and there is a unique hereditary property of partitions with this speed.

**Theorem 4.** Let  $\mathcal{P}$  be a monotone property of ordered graphs. If, for every constant  $c > 0$ , there exists an  $N = N(c) \in \mathbb{N}$  such that  $|\mathcal{P}_N| > c^N$ , then  $\mathcal{P}$  contains the ordered graph  $H(\pi)$  for every  $\pi \in \Pi$ , and hence

$$|\mathcal{P}_n| \geq |\mathcal{P}_n^{(1,2)}| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k! = n^{n/2+o(n)}$$

for every  $n \in \mathbb{N}$ . This lower bound is best possible, and there is a unique monotone property of ordered graphs with this speed.

For  $t \in \mathbb{N}$ , let  $K_t$  denote the complete ordered graph on  $t$  vertices, and let  $K_{t,t}$  denote the complete ordered bipartite graph on  $[2t]$  with edge set  $E(K_{t,t}) = \{\{i, j\} : i \leq t < j\}$ . We shall deduce the following theorem from Theorem 4.

**Theorem 5.** Let  $\mathcal{P}$  be a hereditary property of ordered graphs such that, for some  $t \in \mathbb{N}$ , neither  $K_t$  nor  $K_{t,t}$  is in  $\mathcal{P}$ . If, for every constant  $c > 0$ , there exists an  $N = N(c) \in \mathbb{N}$  such that  $|\mathcal{P}_N| > c^N$ , then  $\mathcal{P}$  contains the ordered graph  $H(\pi)$  for every  $\pi \in \Pi$ , and hence

$$|\mathcal{P}_n| \geq |\mathcal{P}_n^{(1,2)}| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k! = n^{n/2+o(n)}$$

for every  $n \in \mathbb{N}$ . This lower bound is best possible, and there is a unique hereditary property containing neither  $K_t$  nor  $K_{t,t}$  with this speed.

We conjecture that Theorems 1 and 3–5 have the following common generalization.

**Conjecture 1.** Let  $\mathcal{P}$  be a hereditary property of ordered hypergraphs. If, for every constant  $c > 0$ , there exists an  $N = N(c) \in \mathbb{N}$  such that  $|\mathcal{P}_N| > c^N$ , then

$$|\mathcal{P}_n| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k! = n^{n/2+o(n)}$$

for every  $n \in \mathbb{N}$ .

The following statement is a special case of Conjecture 1, but still generalizes Theorems 3–5 (since partitions can be represented by ordered graphs whose components are complete graphs), and would be very interesting in its own right. It was in fact our main motivation for studying ordered hypergraphs and partitions.

**Conjecture 2.** Let  $\mathcal{P}$  be a hereditary property of ordered graphs. If, for every constant  $c > 0$ , there exists an  $N = N(c) \in \mathbb{N}$  such that  $|\mathcal{P}_N| > c^N$ , then

$$|\mathcal{P}_n| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k! = n^{n/2+o(n)}$$

for every  $n \in \mathbb{N}$ .

**Remark 2.** Note that in both conjectures the lower bounds, if true, are the best possible (by Lemma 6, below). However, it is *not* true that, under the conditions of the conjectures,  $\mathcal{P}$  must contain the ordered graph  $H(\pi)$  on  $[2k]$  with edge set  $\{\{i, \pi(i) + k\} : i \in [k]\}$  for every  $k \in \mathbb{N}$  and  $\pi \in \Pi_k$ . To see this, call an ordered graph  $G$  on  $[n]$  *co-matching* if  $\{x_1, y_1\}, \{x_2, y_2\} \in \binom{[n]}{2} \setminus E(G)$  implies that  $|\{x_1, y_1\} \cap \{x_2, y_2\}| = 0$  or  $2$ , and call  $G$  *star-matching* if (say)  $\{x_1, y_1\} \in E(G)$  implies that  $\{x_1, y_2\} \in E(G)$  for every  $y_1 \leq y_2 \leq n$ . The collection of all co-matchings and the collection of all star-matchings are hereditary properties of ordered graphs with super-exponential speeds, but neither contains all the graphs  $H(\pi)$ .

For further details on the possible speeds of hereditary properties of ordered graphs, see [3], which considers such properties with speed below  $2^n$ , and also those with speed above  $2^{\epsilon n^2}$ .

The rest of the paper is organized as follows. In Section 2 we shall state the Klazar–Marcus–Tardos and Marcus–Tardos theorems, and show that the former is implied by each of Theorems 3–5; in Section 3 we shall prove Theorem 2 using the Marcus–Tardos theorem; in Section 4 we shall deduce Theorem 1 from Theorem 2; and in Section 5 we shall deduce Theorems 3–5.

## 2. The Klazar–Marcus–Tardos and Marcus–Tardos theorems

We begin by recalling the theorems of Marcus and Tardos [22].

Given  $n \in \mathbb{N}$ , we shall call a permutation of  $[n]$  an  $n$ -permutation. An  $n$ -permutation  $\pi$  is said to *contain* a  $k$ -permutation  $\sigma$  if there are integers  $1 \leq a(1) < \dots < a(k) \leq n$  such that  $\pi(a(i)) < \pi(a(j))$  if and only if  $\sigma(i) < \sigma(j)$ . Otherwise  $\pi$  is said to *avoid*  $\sigma$ . A property of permutations is a collection of permutations, closed under isomorphism. A property of permutations is said to be hereditary if it is also closed under containment.

The following theorem was conjectured by Stanley and Wilf around 1992 (see [2,10,22]), and proved by Marcus and Tardos in 2004 (Corollary 2 of [22]), using a theorem of Klazar [16]. This result is usually known as the Stanley–Wilf conjecture, but we shall refer to it as the Klazar–Marcus–Tardos theorem, or simply as Theorem A.

**Theorem A.** *Let  $\mathcal{P}$  be a hereditary property of permutations. Either  $\mathcal{P}$  is the set  $\Pi$  of all permutations, so  $|\mathcal{P}_n| = n!$  for every  $n \in \mathbb{N}$ , or there exists a constant  $c = c(\mathcal{P})$  such that  $|\mathcal{P}_n| \leq c^n$  for every  $n \in \mathbb{N}$ .*

We also state here the theorem of Marcus and Tardos (Theorem 1 of [22]), which was originally conjectured by Füredi and Hajnal in [14]. In Section 3 we shall use it to prove Theorem 2.

**Theorem B.** *For every permutation matrix  $M$ , there exists a constant  $C = C(M)$  such that any  $n \times n$  (0, 1)-matrix with at least  $Cn$  of its entries 1 contains  $M$ .*

In this section we shall show that the simplest case of Conjectures 1 and 2 (the case in which every  $G \in \mathcal{P}$  is an ordered graph with maximum degree of at most one) is equivalent to the Klazar–Marcus–Tardos theorem, and deduce that our main results generalize that theorem. We start, however, by proving the following lemma, which gives the final implication of Theorems 1 and 3–5.

**Lemma 6.** Let  $\mathcal{P}$  be a hereditary property of ordered hypergraphs. If  $H(\pi) \in \mathcal{P}$  for every  $\pi \in \Pi$ , then

$$|\mathcal{P}_n| \geq |\mathcal{P}_n^{(1,2)}| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!$$

for every  $n \in \mathbb{N}$ . Moreover, there is a unique hereditary property containing every  $H(\pi)$  with this speed.

**Proof.** Given integers  $k, n \in \mathbb{N}$ , a subset  $A \subset [n]$  of size  $2k$  (with elements  $a(1) < \dots < a(2k)$  say), and permutation  $\pi \in \Pi_k$ , define  $G(n, A, \pi)$  to be the ordered hypergraph on vertex set  $[n]$ , and with edge set  $\{\{a(i), a(\pi(i) + k)\} : i \in [k]\}$ . Let  $\mathcal{P}$  be a hereditary property of ordered hypergraphs with  $H(\pi) \in \mathcal{P}$  for every  $\pi \in \Pi$ . We shall show that  $G(n, A, \pi) \in \mathcal{P}$  for every such  $n, A$  and  $\pi$ .

Indeed, let  $n, A$  and  $\pi$  be as described, let  $X$  be the set of isolated vertices in  $G = G(n, A, \pi)$ , let  $Y = \{v \in X : v < a_k\}$  and let  $Z = X \setminus Y$ . Suppose that  $|Y| = r$  and  $|Z| = s$ , and consider an ordered graph  $H$  on  $[-s + 1, n + r]$  formed by adding to  $G$  an arbitrary matching between the vertices  $\{-s + 1, \dots, 0\}$  and  $Z$ , and an arbitrary matching between the vertices  $\{n + 1, \dots, n + r\}$  and  $X$ . It is easy to see that  $H$  is isomorphic to  $H(\sigma)$  for some  $\sigma \in \Pi_{k+r+s}$ , so  $H \in \mathcal{P}$ , and that  $G$  is an induced subgraph of  $H$ , so  $G \in \mathcal{P}$ .

Thus  $\mathcal{P}$  contains the ordered hypergraph  $G(n, A, \pi)$  for every  $k \in \mathbb{N}$ ,  $\pi \in \Pi_k$ ,  $n \geq 2k$  and  $A \subset [n]$  with  $|A| = 2k$ , and hence

$$\begin{aligned} |\mathcal{P}_n| &\geq |\mathcal{P}_n^{(1,2)}| \geq |\{(k, \pi, A) : 2k \leq n, \pi \in \Pi_k \text{ and } A \subset [n]^{(2k)}\}| \\ &= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!. \end{aligned}$$

Finally, note that the collection  $\mathcal{Q} = \{G(n, \pi, A) : \pi \in \Pi_k \text{ and } A \subset [n]^{(2k)} \text{ for some } k \leq n/2\}$  forms a hereditary property of ordered hypergraphs, and  $H(\pi) \in \mathcal{Q}$  for every  $\pi \in \Pi$ . By the argument above, if  $H(\pi) \in \mathcal{P}$  for every  $\pi \in \Pi$  then  $\mathcal{Q} \subset \mathcal{P}$ , so  $\mathcal{Q}$  is the unique such hereditary property of ordered hypergraphs with this speed.  $\square$

We shall now show that the simplest case of [Conjecture 1](#) follows from the Klazar–Marcus–Tardos theorem. We shall not need this to prove our main results, but the proof is short and has some independent value. Here we use  $\mathcal{G}$  to denote a property of ordered graphs, to distinguish it from a property  $\mathcal{P}$  of permutations.

**Theorem 7.** Let  $\mathcal{G}$  be a hereditary property of ordered graphs of maximal degree at most 1. If, for every constant  $c > 0$ , there exists an  $N = N(c) \in \mathbb{N}$  such that  $|\mathcal{G}_N| > c^N$ , then  $\mathcal{G}$  contains the ordered graph  $H(\pi)$  for every  $\pi \in \Pi_k$ , and hence

$$|\mathcal{G}_n| \geq |\mathcal{G}_n^{(1,2)}| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!$$

for every  $n \in \mathbb{N}$ .

**Proof.** Let  $\mathcal{G}$  be a hereditary property of ordered graphs of maximal degree at most one, and suppose that, for every constant  $c_1 > 0$ , there exists an  $N = N(c_1) \in \mathbb{N}$  such that  $|\mathcal{G}_N| > c_1^N$ . Given an ordered graph  $G \in \mathcal{G}$ , we define a permutation  $\phi(G)$ . Suppose that  $G$  has  $k$  edges,

$e_1 = \{a_1, b_1\}, \dots, e_k = \{a_k, b_k\}$  (where  $a_i < b_i$  for each  $i \in [k]$ ), ordered by their left-endpoints, i.e.  $a_i < a_j$ , if and only if  $i < j$  (recall that  $\Delta(G) \leq 1$ ). Let  $\pi$  be the  $k$ -permutation such that  $b_{\pi(1)} < \dots < b_{\pi(k)}$ , and define  $\phi(G) = \pi^{-1}$ .

Let  $\mathcal{P} = \{\phi(G) : G \in \mathcal{G}\}$ . Since  $\mathcal{G}$  is hereditary, so is  $\mathcal{P}$ , since removing a vertex from a permutation corresponds to removing one of the endpoints of the corresponding edge. By the Klazar–Marcus–Tardos theorem, either  $\mathcal{P} = \Pi$  or there exists a constant  $c$  such that  $|\mathcal{P}_n| \leq c^n$  for every  $n \in \mathbb{N}$ .

Suppose the latter, so  $|\mathcal{P}_n| \leq c^n$  for every  $n \in \mathbb{N}$ . Assuming  $c > 2$ , we claim that in this case

$$|\mathcal{G}_n| \leq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{k} \binom{n-k}{k} c^k < n \binom{n}{\lfloor n/2 \rfloor}^2 c^{\lfloor n/2 \rfloor} < (4\sqrt{c})^n.$$

To see this, simply note that any ordered graph  $G$  of maximal degree at most one is determined by its order, its left-endpoint set, its right-endpoint set, and the permutation  $\phi(G)$ . Hence, setting  $c_1 = 4\sqrt{c}$ , we have a contradiction to our assumption that  $|\mathcal{G}_N| > c_1^N$  for some  $N \in \mathbb{N}$ .

Next, suppose that  $\mathcal{P} = \Pi$ . We want to show that  $\mathcal{G}$  contains  $H(\pi)$  for every  $\pi \in \Pi$ , so let us fix  $k \in \mathbb{N}$  and  $\pi \in \Pi_k$ . Let  $\pi'$  be the  $(k+1)$ -permutation defined as follows:  $\pi'(i) = \pi(i) + 1$  for each  $i \in [k]$ , and  $\pi'(k+1) = 1$ . Since  $\mathcal{P} = \Pi$ , we have  $\pi' \in \mathcal{P}$ , so for some  $G \in \mathcal{G}$  we have  $\pi' = \phi(G)$ .

Now notice that in  $G$ , all left-endpoints occur to the left of *all* right-endpoints, since  $\pi'(k+1) = 1$ . Therefore, letting  $G'$  be a subgraph of  $G$  induced by the first  $k$  left-endpoints and last  $k$  right-endpoints, we have  $G' = H(\pi)$ .  $\square$

**Remark 3.** In fact, for any permutation  $\pi \in \Pi_k$ , the number of ordered graphs  $G$  of order  $n$  and maximal degree at most one with  $\phi(G) = \pi$  is at most  $\binom{n}{k} \text{Cat}(k)$ , where  $\text{Cat}(k) = \frac{1}{k+1} \binom{2k}{k}$  is the  $k$ th Catalan number. To see this, we use the fact that there are exactly  $\text{Cat}(k)$  legal sequences of  $k$  left- and  $k$  right-brackets (i.e. in any initial segment of the sequence there are at least as many left-brackets as right-brackets). Given any ordered graph  $G \in \mathcal{G}$ , we can define a corresponding sequence of brackets,  $\psi(G)$ , by taking a left-bracket for every vertex which is the left-endpoint of an edge, and a right-bracket for every right-endpoint.

Now, given  $n$ ,  $\phi = \phi(G)$ ,  $\psi = \psi(G)$  and the (even-sized) subset  $A = \{v \in [n] : d_G(v) = 1\}$ , it is simple to reconstruct  $G$ : if the elements of  $A$  are  $a(1) < \dots < a(2k)$ , the left-brackets of  $\psi$  lie in positions  $1 \leq s(1) < \dots < s(k) \leq 2k$  and the right-brackets lie in positions  $1 \leq t(1) < \dots < t(k) \leq 2k$  (so  $\{s(1), \dots, s(k), t(1), \dots, t(k)\} = [2k]$ ), then the edge set is  $\{\{a(s(i)), a(t(\phi(i)))\} : i \in [k]\}$ . Note that, although for many permutation–bracket sequence pairs  $(\phi, \psi)$  no ordered graph  $G$  has  $\phi(G) = \phi$  and  $\psi(G) = \psi$  (for example,  $\phi = 21$  and  $\psi = ()()$ ), for the identity permutation all  $\text{Cat}(k)$  bracket pairs are realised.

We proved Theorem 7 using Theorem A; we now prove the reverse implication. It will follow almost immediately that Theorem  $i$  implies Theorem A for  $i = 3$  and 4 and 5.

**Lemma 8.** *Theorem 7 implies Theorem A.*

**Proof.** Let  $\mathcal{P}$  be a non-trivial hereditary property of permutations (i.e. different from  $\Pi$ ), and assume that Theorem 7 holds. Let the ordered graphs  $G(n, A, \pi)$  be as defined above, and let

$$\mathcal{G} = \{G(n, A, \pi) : n \in \mathbb{N}, A \subset [n], \pi \in \mathcal{P}, |A| = 2|\pi|\}.$$

Because  $\mathcal{P}$  is hereditary,  $\mathcal{G}$  is also hereditary, since removing an isolated vertex from  $G(n, A, \pi)$  gives  $G(n-1, A', \pi)$  (for some  $A' \subset [n-1]$ ), and removing a non-isolated vertex corresponds to removing an element from  $\pi$ .

Since  $\mathcal{P} \neq \Pi$ , there exists some  $\pi \notin \mathcal{P}$ , and, by definition,  $\mathcal{G}$  does not contain  $H(\pi)$ . Hence, by Theorem 7, there exists a constant  $c > 0$  such that  $|\mathcal{G}_n| \leq c^n$  for every  $n \in \mathbb{N}$ . But  $|\mathcal{G}_{2n}| \geq |\mathcal{P}_n|$  for every  $n \in \mathbb{N}$ , so  $|\mathcal{P}_n| \leq c^{2n}$  for every  $n \in \mathbb{N}$ .  $\square$

We can now deduce that our main theorems do indeed generalize the Klazar–Marcus–Tardos theorem.

**Corollary 9.** *Each of the Theorems 3–5 implies Theorem A.*

**Proof.** To show that Theorems 3 and 5 imply Theorem A, it suffices to observe that any hereditary property of ordered graphs of maximal degree at most one may be viewed as a hereditary property of partitions (with part sizes of at most 2) or as a hereditary property of ordered graphs containing no  $K_3$  and no  $K_{2,2}$ . The result then follows by Lemma 8.

To show that Theorem 4 implies Theorem A, let  $\mathcal{P}$  be a hereditary property of ordered graphs of maximal degree one, and consider the minimal monotone property of ordered graphs  $\mathcal{P}'$  containing  $\mathcal{P}$ . If  $\mathcal{P}'$  contains the ordered graph  $H(\pi)$  (for some  $\pi \in \Pi$ ) then so does  $\mathcal{P}$ . Otherwise  $|\mathcal{P}'_n| \leq c^n$  for some  $c > 0$  and every  $n \in \mathbb{N}$  by Theorem 4, and hence  $|\mathcal{P}_n| \leq |\mathcal{P}'_n| \leq c^n$  for every  $n \in \mathbb{N}$ . The result again follows by Lemma 8.  $\square$

### 3. Proof of the generalized Marcus–Tardos theorem

In this section we shall prove Theorem 2. Recall that, by Theorem B, for each permutation matrix  $M$ , there exists a constant  $C(M)$  such that any  $n \times n(0, 1)$ -matrix with at least  $C(M)n$  of its entries 1 contains  $M$ . For each  $k \in \mathbb{N}$ , let  $C(k)$  be the constant obtained in the Theorem B for  $k \times k$  matrices, i.e.  $C(k) = \max\{C(M) : M \text{ a } k \times k \text{ permutation matrix}\}$ . We shall give our bounds on  $c_k$  in terms of  $C(k)$ .

To obtain Theorem 2, we use Theorem B to prove it in the case that the rows each have a bounded number of 1s, and then use this result and the method of Marcus and Tardos [22] to prove the general case. First, however, we need to show that Theorem B implies Theorem 2 in the case that each row has exactly two 1s; in fact, these statements are equivalent.

**Lemma 10.** *Let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be any function. The following statements satisfy (i)  $\Rightarrow$  (ii)  $\Rightarrow$  (iii).*

- (i) *For each  $k, m, n \in \mathbb{N}$ , any  $m \times n(0, 1)$ -matrix with at least  $f(k)n$  of its entries 1, and at most two 1s in each row, and with each row different, contains a member of each class of  $\mathcal{M}(k)$ .*
- (ii) *Theorem B holds with  $C(M_k) = f(k)$  for each  $k \times k$  permutation matrix  $M_k$ .*
- (iii) *For each  $k, m, n \in \mathbb{N}$ , any  $m \times n(0, 1)$ -matrix with at least  $(2f(k+1)+1)n$  of its entries 1, and at most two 1s in each row, and with each row different, contains a member of each class of  $\mathcal{M}(k)$ .*

**Proof.** First we shall prove that (i) implies (ii). Let  $k, m, n \in \mathbb{N}$ ,  $M = (m_{ij})$  be a  $k \times k$  permutation matrix, and  $A = (a_{ij})$  be an  $n \times n(0, 1)$ -matrix with at least  $f(k)n$  1s, with  $f(k)$  given by (i). We wish to show that  $A$  contains  $M$ . Suppose, without loss of generality, that there are more 1s above the top-left/bottom-right diagonal than below it (otherwise, replace  $A$  and  $M$  by  $A^T$  and  $M^T$ ). Let the number of pairs  $(i, j)$  for which  $i \leq j$  and  $a_{ij} = 1$  be  $m$ , and label them  $e_1, \dots, e_m$  arbitrarily. We define an  $m \times n(0, 1)$ -matrix  $B = (b_{ij})$  with at most two 1s in each

row, and with each row different, by letting  $b_{ij} = 1$  if and only if vertex  $j$  is an endpoint of  $e_i$ . Note that at least  $f(k)n$  of the entries of  $B$  are 1.

Applying (i) to  $B$ , we see that  $B$  must contain a matrix  $(K, L) \sim (I, M)$ , where  $(I, M)$  is the  $k \times 2k$  matrix obtained by putting the identity matrix in front of  $M$  (so  $(I, M)_{ij} = \delta_{ij}$  and  $(I, M)_{i(j+k)} = m_{ij}$  for each  $j \in [k]$ ). Suppose that  $(K, L)$  occurs in columns  $a_1 < \dots < a_k < b_1 < \dots < b_k$  of  $B$ . Then  $M$  occurs in the intersection of the rows  $a_1, \dots, a_k$  and the columns  $b_1, \dots, b_k$  of  $A$ , and so we are done.

The proof that (ii) implies (iii) is similar. Again, let  $k, m, n \in \mathbb{N}$ ,  $M = (m_{ij})$  be a  $k \times k$  permutation matrix, and let  $B = (b_{ij})$  be an  $m \times n(0, 1)$ -matrix with at most two 1s in each row, with each row different, and at least  $(2f(k+1) + 1)n$  of its entries 1. It will suffice to show that  $B$  contains some matrix  $(K, L) \sim (I, M)$ .

We produce from  $B$  an  $n \times n(0, 1)$ -matrix  $A = (a_{ij})$  by letting  $a_{ij} = 1$  if and only if  $i < j$  and  $b_{ri} = b_{rj} = 1$  for some  $r \in [m]$ . Note that at least  $f(k+1)n$  of the entries of  $A$  are 1. Applying (ii) to  $A$ , we see that  $A$  must contain the  $(k+1) \times (k+1)$  matrix  $M'$ , formed by putting  $M$  in the top right-hand corner and a single 1 in the bottom left-hand corner. Thus  $M'_{i(j+1)} = m_{ij}$  for  $i, j \in [k]$ ,  $M'_{(k+1)1} = 1$ , and  $M'_{ij} = 0$  otherwise. Suppose that  $M'$  occurs in rows  $a_1 < \dots < a_{k+1}$  and columns  $b_1 < \dots < b_{k+1}$  of  $A$ , and that the 1s corresponding to 1-entries of  $M$  correspond to rows  $r_1 < \dots < r_k$  of  $B$ . Since  $A_{a_{k+1}b_1} = 1$  and  $A$  is upper triangular,  $a_{k+1} < b_1$ . Therefore some  $(K, L) \sim (I, M)$  occurs in the intersection of the rows  $r_1, \dots, r_k$  and the columns  $a_1, \dots, a_k, b_2, \dots, b_{k+1}$  of  $B$ , and again we are done.  $\square$

Since Theorem B holds with  $C(M) = C(k)$ , we have the following immediate corollary.

**Corollary 11.** *Let  $k \in \mathbb{N}$ . Any  $m \times n(0, 1)$ -matrix with at least  $(2C(k+1) + 1)n$  of its entries 1, at most two 1s in each row, and with each row different, contains a member of each class of  $\mathcal{M}(k)$ .*

To prove the case where the rows have a bounded number of 1s, we shall use the following trivial observation.

**Lemma 12.** *Let  $G$  be a bipartite graph with parts  $A$  and  $B$ . Suppose that  $d(v) \geq 1$  for each  $v \in A$  and  $d(v) \leq m$  for each  $v \in B$ . Then there exists a matching in  $G$  of size of at least  $|A|/m$ .*

**Proof.** If  $d(v) > 1$  for any  $v \in A$ , then remove all but one of the edges joined to  $v$ . We now have a family of stars, each centred in  $B$  and of order at most  $m+1$ . Take one edge from each.  $\square$

For each  $D \in \mathbb{N}$ , define  $g_D : \mathbb{N} \rightarrow \mathbb{N}$  by  $g_D(x) = \sum_{i=0}^{D-1} (i+1) \binom{x}{i}$ . Note that  $g_D(x) < 2D \binom{x}{D-1}$  if  $x > 3D$ . Let  $\mathcal{H}$  be an ordered hypergraph on  $[n]$  in which every edge has size of at most  $D$ . For each vertex  $v \in [n]$ , the 2-degree of  $v$  in  $\mathcal{H}$  is  $d_{\mathcal{H}}^{(2)}(v) = |\{u \in [n] : \{u, v\} \subset E \text{ for some } E \in E(\mathcal{H})\}|$ . Suppose that a vertex of 2-degree  $x$  is removed from  $\mathcal{H}$ ; by how much can  $\|\mathcal{H}\| = \sum_{E \in E(\mathcal{H})} |E|$  decrease? For each  $0 \leq i \leq D-1$ ,  $v$  is contained in at most  $\binom{x}{i}$  edges  $E$  of  $\mathcal{H}$  of size  $i+1$ , and each of these must be removed entirely if  $E \setminus v$  is also an edge. Hence the maximum possible decrease in  $\|\mathcal{H}\|$  is  $g_D(x)$ .

For each  $k \in \mathbb{N}$ , let  $C_1(k) = 2C(k+1) + 1$ .

**Lemma 13.** *Let  $k, m, n, D \in \mathbb{N}$ , with  $D \geq 2$ , and let  $g_D : \mathbb{N} \rightarrow \mathbb{N}$  be as defined above. Any  $m \times n(0, 1)$ -matrix with at least  $g_D(D(D-1)C_1(k))n$  of its entries 1, at most  $D$  of the entries of each row 1, and with each row different, contains a member of each class of  $\mathcal{M}(k)$ .*

**Proof.** We shall use [Corollary 11](#). Let  $k, m, n, D \in \mathbb{N}$ ,  $M = (m_{ij})$  be a  $k \times k$  permutation matrix, and  $A = (a_{ij})$  be an  $m \times n(0, 1)$ -matrix with at least  $g_D(D(D-1)C_1(k))n$  of its entries 1, at most  $D$  of the entries in each row 1, and each row different. We shall show that  $A$  contains a matrix in the equivalence class of  $(I, M)$ .

Consider the ordered hypergraph  $\mathcal{H}$  on vertex set  $[n]$  with edge set  $\{E_i : i \in [m] \text{ and } a_{ij} = 1 \Leftrightarrow j \in E_i\}$ , so the rows of  $A$  are the indicator functions of the edges. Note that  $\|\mathcal{H}\| = \sum_i |E_i| \geq g_D(D(D-1)C_1(k))n$ . We first wish to find a subset  $S$  of  $[n]$  in which there are at least  $\binom{D}{2} C_1(k)|S|$  distinct pairs  $\{i, j\}$ , each contained in some edge of  $\mathcal{H}$ . If  $d_{\mathcal{H}}^{(2)}(v) < D(D-1)C_1(k)$  for some vertex  $v$  then, by the comments above, removing  $v$  from the ordered hypergraph causes  $\|\mathcal{H}\|$  to decrease by at most  $g_D(D(D-1)C_1(k)-1) < g_D(D(D-1)C_1(k))$ . Thus removing  $v$  causes the density of edges in the ordered hypergraph to increase.

Thus, if we repeatedly remove vertices of minimal 2-degree from  $\mathcal{H}$ , we must eventually produce an ordered hypergraph  $\mathcal{H}'$  on vertex set  $S$  in which every vertex  $v$  has  $d_{\mathcal{H}'}^{(2)}(v) \geq D(D-1)C_1(k)$ . By counting degrees, there are at least  $\binom{D}{2} C_1(k)|S|$  distinct pairs  $\{i, j\} \subset S$ , each contained in some edge of  $\mathcal{H}'$ , and hence of  $\mathcal{H}$ .

Let  $\mathcal{T} = \{\{i, j\} \subset S : i, j \in E \text{ for some } E \in \mathcal{H}\}$  be the set of such pairs. Let  $B$  be the bipartite graph on sets  $\mathcal{T}$  and  $E(\mathcal{H})$  with edges corresponding to containment (i.e.  $(\{i, j\}, E)$  is an edge of  $B$  iff  $\{i, j\} \subset E$ ). By [Lemma 12](#), there exists a matching  $W$  in  $B$  of size  $t$ , with  $t \geq |\mathcal{T}| / \binom{D}{2} \geq C_1(k)|S|$ , since each edge has order at most  $D$ . Let  $\mathcal{T}_1$  be the set of endpoints of  $W$  lying in  $\mathcal{T}$ .

Let  $s = |S|$ , and define  $P$  to be the  $t \times s(0, 1)$ -matrix in which the columns correspond to elements of  $S$ , and the rows are the indicator functions of the edges in  $\mathcal{T}_1$ . In  $P$ , all rows have exactly two 1s, all rows are different, and  $2t \geq 2C_1(k)s$  of its entries are 1, so, by [Corollary 11](#),  $P$  contains some matrix  $(K, L)$  equivalent to  $(I, M)$ .

Fix a copy of  $(K, L)$  in  $P$ , and let the columns of  $P$  containing this  $(K, L)$  be those corresponding to the vertices  $b(1) < \dots < b(2k)$  of  $\mathcal{H}$  (note that here  $b(i)$  is the *original* labelling of the vertex in  $[n]$ ). We claim that the corresponding columns of  $A$  contain a matrix equivalent to  $(K, L)$ . To see this, let the rows of  $P$  containing the same copy of  $(K, L)$  be  $a(1) < \dots < a(k)$ , and let  $p(i)$  be the pair in  $\mathcal{T}_1$  corresponding to row  $a(i)$  for  $1 \leq i \leq k$ . For each of these pairs  $p(i)$ , choose the edge  $e(i) \in E(\mathcal{H})$  it was matched to by  $W$ . We have thus found  $k$  distinct edges  $e(1), \dots, e(k)$ , for which  $p(i) \subset e(i)$ . It follows immediately that the columns  $b(1), \dots, b(2k)$  of  $A$  contain some matrix  $(K', L') \sim (K, L) \sim (I, M)$ . This completes the proof.  $\square$

We are now ready to prove [Theorem 2](#).

**Proof of Theorem 2.** For each  $n, k \in \mathbb{N}$ , let  $f(n, k)$  be the largest number of 1s possible in an  $m \times n(0, 1)$ -matrix  $A$  (where  $m \in \mathbb{N}$  is arbitrary), with each row different, not containing any member of some class of  $\mathcal{M}(k)$ . We wish to show that  $f(n, k) = O(n)$ , where  $k$  is fixed and  $n \rightarrow \infty$ . The proof (that  $f(n, k) < c_k n$ , where  $c_k$  will be determined later) will be by induction on  $n$ . Note that, since each row of  $A$  is different, at most  $n2^n$  of the entries of  $A$  can be 1s. We shall choose  $c_k > 2^{8k^3}$ , so the statement  $f(n, k) < c_k n$  is (trivially) true for  $n \leq 8k^3$ .

Let  $k, m, n \in \mathbb{N}$  with  $n \geq 8k^3$ ,  $M = (m_{ij})$  be a  $k \times k$  permutation matrix, and  $A$  be an  $m \times n(0, 1)$ -matrix, with each row different, not containing any matrix equivalent to  $(I, M)$ . Following the method of Marcus and Tardos, we want to divide  $A$  up into ‘fat’ and ‘skinny’ blocks of size  $1 \times t$  for some  $t$ . In preparation for this, we must remove the rows with few 1s. Let

$t \in \mathbb{N}$  and let  $D = (2k - 1)t$ . (We shall eventually set  $t = 2k^2$ , but we postpone choosing this value until it is clear why the choice is being made. Our argument up to that point works for any  $t \in \mathbb{N}$ .) By Lemma 13, there are at most  $g_D(D - 1)C_1(k)n$  1s in rows with at most  $D$  of their entries 1, otherwise  $A$  would contain some  $(K, L) \sim (I, M)$ , contradicting our assumption. Let  $A' = (a'_{ij})$  be the  $m' \times n$  matrix obtained from  $A$  by deleting the rows with at most  $D$  entries 1.

Now, let  $q$  and  $r$  satisfy  $n = qt + r$ , with  $q \in \mathbb{N}$  and  $r \in [t]$ , and partition  $A'$  into  $qm'$  blocks of size  $1 \times t$  and  $m'$  blocks of size  $1 \times r$  as follows. Let the  $1 \times t$  blocks be  $S_{ij} = (a'_{i\ell} : \ell \in [(j - 1)t + 1, jt])$ , for each  $i \in [m']$  and  $j \in [q]$ , and the  $1 \times r$  blocks be  $S_{i(q+1)} = (a'_{i\ell} : \ell \in [qt + 1, n])$ , for each  $i \in [m']$ . Define  $B = (b_{ij})$  to be the  $m' \times (q + 1)(0, 1)$ -matrix obtained by assigning the value 1 to a block if any entry of the block is 1. Thus, for  $j \in [q]$ ,  $b_{ij} = 0$  if and only if  $a'_{i\ell} = 0$  for every  $\ell \in [(j - 1)t + 1, jt]$ , and similarly for  $j = q + 1$ .

**Claim 1:**  $B$  contains no matrix equivalent to  $(I, M)$ .

**Proof.** This is Lemma 4 of Marcus and Tardos [22]. To spell it out, assume that  $B$  contains such a matrix  $P$  and, for each 1 that it contains, choose an arbitrary non-zero entry from the corresponding blocks of  $A'$ . These represent a copy of  $P$  in  $A$ , a contradiction.  $\square$

Call a block ‘fat’ if at least  $2k$  of its entries are 1.

**Claim 2:** There are at most  $\binom{t}{2k} (k - 1)$  fat blocks in any column of blocks  $(S_{ij} : i \in [m'])$ .

**Proof.** This is Lemma 5 of [22]. If there are more than  $\binom{t}{2k} (k - 1)$  fat blocks in a given column  $S_{ij}$ , then there are at least  $k$  fat blocks which contain 1s in the same  $2k$  columns of  $A'$ . Hence  $A'$  contains a complete  $k \times 2k$  matrix (i.e. a matrix in which all entries are 1), so  $A$  contains every  $k \times 2k(0, 1)$ -matrix, another contradiction.  $\square$

We wish to bound the number of 1s in  $B$ .  $B$  may contain repeated rows but, since every row in  $A'$  has at least  $(2k - 1)t + 1$  of its entries 1, every row of  $B$  must have at least  $2k$  of its entries 1. Thus, if any row occurs in  $B$  more than  $k - 1$  times, then  $B$  contains a complete  $k \times 2k$  matrix, contradicting Claim 1. If we let  $B'$  be the matrix obtained from  $B$  by deleting repeated rows, then  $B'$  contains no matrix equivalent to  $(I, M)$  and all rows of  $B'$  are different, so at most  $f(\lceil n/t \rceil, k)$  of the entries of  $B'$  are 1. Since each row in  $B$  was repeated at most  $(k - 1)$  times, it follows that at most  $(k - 1)f(\lceil n/t \rceil, k)$  of the entries of  $B$  are 1. We have thus established the following recurrence:

$$f(n, k) \leq (2k - 1)kf\left(\left\lceil \frac{n}{t} \right\rceil, k\right) + \binom{t}{2k} (k - 1)n + g_D\left(2\binom{D}{2} C_1(k)\right)n.$$

Let  $t = 2k^2$ . Assuming (by induction) that

$$f\left(\left\lceil \frac{n}{2k^2} \right\rceil, k\right) < c_k \left\lceil \frac{n}{2k^2} \right\rceil,$$

and, using the inequality  $g_D(x) < 2D\binom{x}{D-1}$  noted earlier, we obtain

$$\begin{aligned} f(n, k) &< (2k^2 - k)c_k \left\lceil \frac{n}{2k^2} \right\rceil + \left( \binom{2k^2}{2k} k + 8k^3 \left( 2 \binom{4k^3 - 2k^2}{2} C_1(k) \right) \right) n \\ &< c_k n - \frac{c_k n}{2k} + (2k^2 - k)c_k + 8k^3 \left( \frac{16k^6 C_1(k)}{4k^3} \right) n. \end{aligned} \quad (1)$$

Now if

$$c_k \geq 32k^4 \binom{16k^6 C_1(k)}{4k^3},$$

then

$$\frac{c_k n}{2k} \geq \frac{c_k n}{4k} + 8k^3 \binom{16k^6 C_1(k)}{4k^3} \geq 2k^2 c_k + 8k^3 \binom{16k^6 C_1(k)}{4k^3}, \quad (2)$$

since  $n \geq 8k^3$ . Set  $c_k = 32k^4 \binom{16k^6 C_1(k)}{4k^3}$ , and note that  $c_k > 2^{8k^3}$ . Now inequalities (1) and (2) imply that  $f(n, k) < c_k n$ , so the induction step is complete.  $\square$

**Remark 4.** Marcus and Tardos proved that  $C(k) < 2k^4 \binom{k^2}{k}$ , so the explicit bound that we obtain is roughly  $c_k = O(k^{9k^4})$ .

#### 4. Ordered hypergraphs

We now deduce Theorem 1 from Theorem 2. This implication may be read out of a proof of Klazar (Theorem 2.5 of [18]) but, for the sake of completeness, we shall prove it (and in fact our proof is slightly different from that in [18]).

Given an ordered hypergraph  $\mathcal{H}$  on  $[n]$ , say that  $\mathcal{H}$  contains a  $k$ -permutation  $\pi$  if  $\mathcal{H}$  contains the ordered hypergraph  $H(\pi)$ . In other words, there exist  $2k$  vertices  $v_1, \dots, v_{2k} \in [n]$  and  $k$  distinct edges  $E_1, \dots, E_k \in E(\mathcal{H})$  such that, letting  $e_i = \{v_i, v_{\pi(i)+k}\}$  denote the edges of  $H(\pi)$ , we have  $e_i \subset E_i$  for each  $i \in [k]$ . Otherwise, say that  $\mathcal{H}$  avoids  $\pi$ . For each permutation  $\pi$ , and each  $n \in \mathbb{N}$ , let  $T_n(\pi)$  denote the family of ordered hypergraphs on  $[n]$  avoiding  $\pi$ .

**Lemma 14.** Let  $k \in \mathbb{N}$  and  $\pi \in \Pi_k$ . If  $\mathcal{H} \in T_n(\pi)$ , then

$$\|\mathcal{H}\| = \sum_{E \in E(\mathcal{H})} |E| < c_k n.$$

**Proof.** The lemma is a simple corollary of Theorem 2. To spell it out, let  $k \in \mathbb{N}$ ,  $\pi \in \Pi_k$  and  $\mathcal{H} \in T_n(\pi)$ , let  $m = |E(\mathcal{H})|$ , and define  $A$  to be an  $m \times n(0, 1)$ -matrix whose rows are the indicator functions of the edges of  $\mathcal{H}$ . (Note that  $A$  is unique up to permutations of its rows.)

Now,  $A$  has exactly  $\|\mathcal{H}\|$  of its entries 1, and each of its rows are different, so, by Theorem 2, if  $\|\mathcal{H}\| \geq c_k n$  then  $A$  contains a sub-matrix  $B \sim (I, M)$ , where  $M$  is the permutation matrix of  $\pi$ . Now, let  $v_1, \dots, v_{2k}$  be the vertices in  $[n]$  corresponding to the columns of  $B$ , and  $E_1, \dots, E_k$  be the edges of  $\mathcal{H}$  corresponding to the rows of  $B$ , ordered so that  $v_i \in E_i$  for each  $i \in [k]$ . Then, for each  $i \in [k]$ , we have  $v_i, v_{\pi(i)+k} \in E_i$ , so  $\mathcal{H}$  contains  $\pi$ , a contradiction. Hence  $\|\mathcal{H}\| < c_k n$ .  $\square$

**Proof of Theorem 1.** Let  $n, k \in \mathbb{N}$ , and  $\pi \in \Pi_k$ . We claim that

$$|T_{2n}(\pi)| \leq |T_n(\pi)| 3^{2c^2 n}, \quad (3)$$

where  $c = c_k$  is the constant obtained in Theorem 2.

To prove inequality (3), we map each  $\mathcal{H} \in T_{2n}(\pi)$  to the ordered hypergraph  $\mathcal{K}$  on  $[n]$  with edge set

$$\{E \subset [n] : \exists E' \in E(\mathcal{H}) \text{ with } i \in E \Leftrightarrow \{2i-1, 2i\} \cap E' \neq \emptyset\}.$$

In other words,  $\mathcal{K}$  is formed by identifying vertices  $2i-1$  and  $2i$  for every  $i \in [n]$ .

We claim that  $\mathcal{K} \in T_n(\pi)$ , i.e. that  $\mathcal{K}$  avoids  $\pi$ . Indeed, suppose for a contradiction that there exist edges  $E_1, \dots, E_k \in E(\mathcal{K})$  and vertices  $v_1, \dots, v_{2k} \in [n]$  such that, for each  $i \in [k]$ , we have  $e_i = \{v_i, v_{\pi(i)+k}\} \subset E_i$ . For each edge  $E_i$ , choose an edge  $F_i$  of  $\mathcal{H}$  such that  $\{2j-1, 2j\} \cap F_i \neq \emptyset$  if and only if  $j \in E_i$  (such an  $F_i$  exists by the definition of  $\mathcal{K}$ ). The edges  $F_i$  are distinct (since the edges  $E_i$  are), and  $e_i \subset F_i$  for each  $i \in [k]$ , so  $\mathcal{H}$  contains  $\pi$ , which is the desired contradiction.

Now, how many ordered hypergraphs  $\mathcal{H}$  map to the same ordered hypergraph  $\mathcal{K}$ ? An edge  $E$  of  $\mathcal{K}$  is the image of  $3^{|E|}$  different possible edges of  $\mathcal{H}$ , since each vertex  $v$  of  $E$  may have come from  $\{2v-1\}$ ,  $\{2v\}$  or  $\{2v-1, 2v\}$ . However, suppose that at least  $2c = 2c_k$  of these did in fact occur in  $\mathcal{H}$  for a given edge  $E$ . Each such edge has a size of at least  $|E|$ , so the ordered hypergraph  $\mathcal{H}'$  induced by  $\mathcal{H}$  on vertex set  $\{2i-1, 2i \in [2n] : i \in E\}$  has  $\|\mathcal{H}'\| \geq 2c|E|$ . But now  $\mathcal{H}' \notin T_{2|E|}(\pi)$ , by Lemma 14, so  $\mathcal{H}$  contains  $\pi$ , a contradiction.

Thus, for each edge  $E$  of  $\mathcal{K}$ , at most  $2c-1$  of the edges that map to it actually occur in  $\mathcal{H}$ , so we have at most

$$\sum_{i=0}^{2c-1} \binom{3^{|E|}}{i} < 3^{2c|E|}$$

choices for these edges.

Now, again by Lemma 14, since  $\mathcal{K} \in T_n(\pi)$ , we have  $\|\mathcal{K}\| \leq cn$ . Thus the maximum possible number of ordered hypergraphs  $\mathcal{H}$  that map to a given  $\mathcal{K}$  is

$$\prod_{E \in E(\mathcal{K})} 3^{2c|E|} = 3^{2c \sum_E |E|} \leq 3^{2c^2 n}.$$

This proves inequality (3).

Now, let  $\mathcal{P}$  be a strongly monotone property of ordered hypergraphs, let  $k \in \mathbb{N}$ ,  $\pi \in \Pi_k$ , and suppose that  $H(\pi) \notin \mathcal{P}$ . Then  $\mathcal{P}_n \subset T_n(\pi)$  and, by inequality (3) and induction on  $n$ ,

$$T_n(\pi) \leq 3^{3c^2 n},$$

for every  $n \in \mathbb{N}$ , where again  $c = c_k$ , since

$$|T_{2n-1}(\pi)| \leq |T_{2n}(\pi)| \leq |T_n| 3^{2c^2 n} \leq 3^{5c^2 n} \leq 3^{3c^2 (2n-1)}$$

when  $n \geq 3$ . So  $|\mathcal{P}_n| \leq 3^{3c^2 n}$  for every  $n \in \mathbb{N}$ , which proves Theorem 1.  $\square$

## 5. Partitions and ordered graphs

We shall now deduce Theorems 3–5 from Theorem 1. We begin with Theorem 3. The implication is very simple, but in any case we shall write out all the details.

**Proof of Theorem 3.** Let  $\mathcal{P}$  be a hereditary property of partitions. For each partition  $P \in \mathcal{P}$ , let  $\mathcal{H}(P)$  be the ordered hypergraph whose edges are the parts of  $P$  of size of at least two. To be precise, if  $P$  is the partition  $\{A_1, \dots, A_t\}$  of  $[n]$ , then  $\mathcal{H}(P)$  has vertex set  $[n]$  and edge set  $\{A_i : i \in [t], |A_i| \geq 2\}$ . Observe that, for any permutation  $\pi$ ,  $\mathcal{H}(P)$  contains  $\pi$  if and only if  $P$  contains the partition  $H(\pi)$  as an induced subpartition.

Suppose that, for some  $\pi \in \Pi$ ,  $\mathcal{P}$  does not contain  $H(\pi)$ . By the observation above,  $\mathcal{H}(P)$  avoids  $\pi$  for every  $P \in \mathcal{P}$ . Now, let  $T(\pi)$  denote the strongly monotone property of ordered hypergraphs consisting of all ordered hypergraphs avoiding  $\pi$  (so  $T(\pi) = \bigcup_n T_n(\pi)$ , with  $T_n(\pi)$  as in the previous section). Then  $\mathcal{H}(P) \in T(\pi)$  for every  $P \in \mathcal{P}$ .

Now apply [Theorem 1](#) to  $T(\pi)$ . Since  $H(\pi) \notin T(\pi)$ , there exists a constant  $c$  such that  $|T_n(\pi)| \leq c^n$  for every  $n \in \mathbb{N}$ . But now we are done, since

$$|\mathcal{P}_n| = |\{\mathcal{H}(P) : P \in \mathcal{P}_n\}| \leq |T_n(\pi)| \leq c^n$$

for every  $n \in \mathbb{N}$ , since  $\mathcal{H}(P) \in T(\pi)$  for every  $P \in \mathcal{P}$ . This proves [Theorem 3](#).  $\square$

**Remark 5.** If  $\pi \in \Pi$ , let  $\mathcal{P}(\pi)$  be the largest hereditary property of partitions such that  $\mathcal{P}$  avoids the partition  $H(\pi)$ . Let  $c'_k$  be the smallest constant such that  $|\mathcal{P}(\pi)_n| < (c'_k)^n$  for every  $n \in \mathbb{N}$  and  $\pi \in \Pi_k$ . We have shown that  $c'_k = O(3^{k^{19k^4}})$ .

We next deduce [Theorem 4](#) from [Theorem 1](#). Since a monotone property of ordered graphs is a strongly monotone property of ordered hypergraphs, the implication is trivial.

**First proof of Theorem 4.** Let  $\mathcal{P}$  be a monotone property of ordered graphs and, for each  $G \in \mathcal{P}$ , let  $G'$  be the ordered hypergraph with the same vertex and edge set as  $G$ . Define  $\mathcal{P}' = \{G' : G \in \mathcal{P}\}$ . Now  $\mathcal{P}'$  is a strongly monotone property of ordered hypergraphs, since each edge of  $G' \in \mathcal{P}'$  has size 2, so the only ordered hypergraphs contained in  $G'$  are its subgraphs. The result now follows by applying [Theorem 1](#) to  $\mathcal{P}'$ .  $\square$

In fact, one can also prove [Theorem 4](#) without using [Theorem 1](#), but using the Marcus–Tardos and Klazar–Marcus–Tardos theorems instead. Alternative proofs can often give new insight into the difficulties and the true nature of a problem, as well as offering new methods for generalization. For this reason, we give a sketch of this second proof.

**Sketch of the second proof of Theorem 4.** Let  $\mathcal{P}$  be a monotone property of ordered graphs, let  $\pi \in \Pi$ , and suppose that  $\mathcal{P}$  does not contain the ordered graph  $H(\pi)$ .

Suppose first that, for some  $n \in \mathbb{N}$ , there exists an ordered graph  $G \in \mathcal{P}_n$  with at least  $C(k+1)n$  edges (where  $C(k)$  is again the constant in [Theorem B](#)). In this case, we can use [Theorem B](#) to find  $H(\pi)$  in  $G$ , just as in the proof of [Lemma 10](#). So assume that, for every  $G \in \mathcal{P}$ ,  $e(G) < C(k+1)|G|$ .

Let  $\mathcal{S}(n, m)$  denote the family of sequences  $(a_1, \dots, a_n)$  such that  $a_i \in \mathbb{N} \cup \{0\}$  for each  $i$  and  $\sum_i a_i = m$ , and let  $\mathcal{S} = \bigcup_{n,m} \mathcal{S}(n, m)$ . We define a map  $\varphi : \mathcal{P} \rightarrow \Pi \times \mathcal{S} \times \mathcal{S}$  as follows.

Let  $G \in \mathcal{P}$  have  $n$  vertices and  $m$  edges. We put the following two linear orders,  $<_\ell$  and  $<_r$ , on the edges of  $G$ . If  $e = \{e_1, e_2\}$  and  $f = \{f_1, f_2\}$  with  $e_1 < e_2$  and  $f_1 < f_2$ , then  $e <_\ell f$  if  $e_1 < f_1$  or  $e_1 = f_1$  and  $e_2 < f_2$ , while  $e <_r f$  if  $e_2 < f_2$  or  $e_2 = f_2$  and  $e_1 < f_1$ . Let  $\varphi_p(G)$  be the  $m$ -permutation that takes the order of the edges under  $<_r$  to the order under  $<_\ell$ . Let  $\varphi_\ell(G)$  be the left-endpoint degree sequence of  $G$ , i.e. the sequence  $(a_1, \dots, a_n)$  where  $a_i$  is the number of edges of  $G$  whose left-endpoint is vertex  $i$ , and let  $\varphi_r(G)$  be the right-endpoint degree sequence of  $G$ . Let  $\varphi(G) = (\varphi_p(G), \varphi_\ell(G), \varphi_r(G)) \in \Pi_n \times \mathcal{S}(n, m) \times \mathcal{S}(n, m)$ .

Let  $\mathcal{Q} = \{\varphi_p(G) : G \in \mathcal{P}\}$ .  $\mathcal{Q}$  is a hereditary property of permutations so, by the Klazar–Marcus–Tardos theorem, either  $\mathcal{Q} = \Pi$  or there exists a constant  $c$  such that  $|\mathcal{Q}_n| \leq c^n$  for every  $n \in \mathbb{N}$ .

Suppose first that  $|\mathcal{Q}_n| \leq c^n$  for every  $n \in \mathbb{N}$ . We claim that, for any  $\varphi \in \Pi \times \mathcal{S} \times \mathcal{S}$ , there is only at most one  $G$  such that  $\varphi(G) = \varphi$ . We omit the proof, which is by induction on  $m$ . For the induction step, remove the first edge of  $G$  in the order  $<_r$ .

So  $|\mathcal{P}_n|$  is just  $|\text{Im}_\varphi(\mathcal{P}_n)|$ , which can easily be approximated, since each  $G \in \mathcal{P}_n$  has at most  $C(k+1)n$  edges. Thus

$$|\text{Im}_{\varphi_p}(\mathcal{P}_n)| \leq \left| \bigcup_{m=0}^{C(k+1)n} \mathcal{Q}_m \right|$$

$$\leq \sum_{k=0}^{C(k+1)n} c^m < 2c^{C(k+1)n},$$

assuming (as we may) that  $c > 2$ , and

$$|\text{Im}_{\varphi_\ell}(\mathcal{P}_n)| \leq \binom{m+n-1}{n-1} < \binom{(C(k+1)+1)n}{n-1} < 2^{(C(k+1)+1)n},$$

and similarly for  $|\text{Im}_{\varphi_r}(\mathcal{P}_n)|$ . Hence

$$\begin{aligned} |\mathcal{P}_n| = |\text{Im}_{\varphi}(\mathcal{P}_n)| &\leq |\text{Im}_{\varphi_p}(\mathcal{P}_n)| \cdot |\text{Im}_{\varphi_\ell}(\mathcal{P}_n)| \cdot |\text{Im}_{\varphi_r}(\mathcal{P}_n)| \\ &< 2^{2(C(k+1)+1)n+1} c^{C(k+1)n}, \end{aligned}$$

so we are done in this case.

Now suppose that  $\mathcal{Q} = \Pi$ . We want to show that  $\mathcal{P}$  contains the ordered graph  $H(\pi)$ , and thus obtain a contradiction. To do this, define the  $2k$ -permutation  $\sigma$  by  $\sigma(2i-1) = 2\pi(i)$  and  $\sigma(2i) = 2\pi(i)-1$  for  $1 \leq i \leq k$  (so, for example, if  $\pi = 213$  then  $\sigma = 432165$ ). By assumption, there exists an ordered graph  $G \in \mathcal{P}$  such that  $\varphi_p(G) = \sigma$ . Let the edges of  $G$  be  $e_1, \dots, e_{2k}$  in the order  $<_\ell$ . Note that, for each  $i \in [k]$ , the edges  $e_{2i-1}$  and  $e_{2i}$  do not share an endpoint, since  $e_{2i-1} <_\ell e_{2i}$  and  $e_{2i-1} >_r e_{2i}$ .

Consider the edges  $\{e_{2i-1} : i \in [k]\}$ . Suppose that two of them share a left (right) endpoint  $v$ . Then all edges between them in the order  $<_\ell$  ( $>_r$ ) also share that endpoint. This contradicts the previous observation that  $e_{2i-1}$  and  $e_{2i}$  do not share an endpoint, so these edges are in fact independent. Thus there exists  $G \in \mathcal{P}$  with  $\Delta(G) = 1$  and  $\varphi_p(G) = \pi$ .

Now apply the same technique to the  $(k+1)$ -permutation  $\pi'$ , where  $\pi'(i) = \pi(i)+1$  if  $i \in [k]$  and  $\pi'(k+1) = 1$ . We obtain  $G \in \mathcal{P}$  with  $\Delta(G) = 1$  and  $\varphi_p(G) = \pi'$ . Again (as in the proof of Theorem 7), notice that, in  $G$ , all left-endpoints occur to the left of *all* right-endpoints, since  $\pi'(k+1) = 1$ . Therefore, letting  $G'$  be a subgraph of  $G$  induced by the first  $k$  left-endpoints and last  $k$  right-endpoints, we have  $G' = H(\pi)$ .

So  $H(\pi) \in \mathcal{P}$ , and this gives us the desired contradiction.  $\square$

Finally, we deduce Theorem 5 from Theorem 4.

**Proof.** Let  $t \in \mathbb{N}$ , and let  $\mathcal{P}$  be a hereditary property of ordered graphs such that  $K_t \notin \mathcal{P}$  and  $K_{t,t} \notin \mathcal{P}$ . Let  $\mathcal{G} = \mathcal{G}(\mathcal{P})$  be the smallest monotone property containing  $\mathcal{P}$ . Note that (trivially)  $|\mathcal{G}_n| \geq |\mathcal{P}_n|$ .

By Theorem 4, either  $\mathcal{G}$  contains  $H(\pi)$  for every  $\pi \in \Pi$ , or there exists a constant  $c$  such that  $|\mathcal{G}_n| < c^n$  for every  $n \in \mathbb{N}$ . Suppose the latter. Then  $|\mathcal{P}_n| \leq |\mathcal{G}_n| < c^n$  for every  $n \in \mathbb{N}$ , in which case we are done. So assume that  $\mathcal{G}$  contains  $H(\pi)$  for every  $\pi \in \Pi$ .

We wish to find, for each permutation  $\pi$ , a large permutation  $\sigma$  such that an ordered graph on  $[n]$  containing no induced copy of  $K_t$  or  $K_{t,t}$ , and containing  $H(\sigma)$  as a subgraph, contains  $H(\pi)$  as an induced subgraph. We shall use Ramsey's theorem and the pigeonhole principle, so recall that, for  $\ell \in \mathbb{N}$ ,  $R(\ell)$  denotes the smallest integer  $n$  such that any graph on  $n$  vertices contains either a clique on  $\ell$  vertices, or an independent set of order  $\ell$ . We shall also write  $S(\ell)$  for the smallest integer  $n$  such that any bipartite graph with one part of order at least  $2\ell-1$  and the other of order at least  $n$  contains either the complete bipartite graph  $K_{\ell,\ell}$  or the empty bipartite graph  $E_{\ell,\ell}$ . It is easy to show that  $R(\ell) < 4^\ell$  and  $S(\ell) < \ell 4^\ell$ . For each  $j \in \mathbb{N}$ , let  $R^{(j+1)}(\ell) = R(R^{(j)}(\ell))$ , where  $R^{(1)}(\ell) = R(\ell)$ , and let  $S^{(j)}(\ell)$  be defined similarly.

Let  $k \in \mathbb{N}$  and  $\pi \in \Pi_k$ . We shall show that  $H(\pi) \in \mathcal{P}$ . Let  $m' = S^{(K)}(t)$ , where  $K = \binom{2k}{2} - k$ , and  $m = R^{(2)}(m')$ . We define the  $mk$ -permutation  $\sigma$  as follows: for each

$i \in [k]$  and  $j \in [0, m-1]$ , let  $\sigma(im-j) = \pi(i)m-j$ . By assumption,  $H(\sigma) \in \mathcal{G}$ ; let  $G \in \mathcal{P}$  be an ordered graph on  $[2mk]$  containing  $H(\sigma)$ . Such a  $G$  must exist by the definition of  $\mathcal{G}$ . We know that, for each  $i \in [mk]$ , the edge  $\{i, \sigma(i)+mk\}$  is in  $E(G)$ . We want to show that, for some subset  $A = \{a(1), \dots, a(2k)\} \subset [2mk]$  with  $a(i) \in [(i-1)m+1, im]$  for each  $i \in [2k]$  and  $a(\pi(i)+k) = a(i) + (k + \pi(i) - i)m$  for each  $i \in [k]$ , these are the only edges induced by  $A$ .

First we use Ramsey's theorem to find 'matching' independent subsets of  $[(i-1)m+1, im]$  for each  $i \in [2k]$ . (This step is necessary if we are to assume only that  $K_t$  and  $K_{t,t}$  are avoided; without it we would have to assume that every ordered graph containing  $K_{t,t}$  is missing from  $\mathcal{P}$ .) We do this first for each  $i \in [k]$ . Since  $m \geq R^{(2)}(m')$ , by Ramsey's theorem there exists either a clique or an independent set  $A_i^{(1)}$  of order  $R(m')$  in  $[(i-1)m+1, im]$ . Since  $R(m') \geq t$ ,  $K_t \notin \mathcal{P}$  and  $\mathcal{P}$  is hereditary,  $A_i^{(1)}$  must be an independent set. For  $i \in [k+1, 2k]$ , let  $A_i^{(1)}$  be the set  $A_{\pi^{-1}(i-k)}^{(1)} + (k + \pi(i) - i)m$  (if  $A$  is a set and  $b \in \mathbb{N}$ , then  $A + b = \{a + b : a \in A\}$ ). Now, again by Ramsey's theorem, for each  $i \in [k+1, 2k]$  there exists a clique or independent set  $A_i^{(2)} \subset A_i^{(1)}$  of order  $m'$ . Again, this must be an independent set, since  $m' \geq t$ . For each  $i \in [k]$ , let  $A_i^{(2)}$  be the set  $A_{\pi(i)+k}^{(2)} - (k + \pi(i) - i)m$ .

Thus we have found independent sets  $A_i^{(2)} \subset [(i-1)m+1, im]$  of order  $m'$  for each  $i \in [2k]$  such that there is a matching between  $A_i^{(2)}$  and  $A_{\pi(i)+k}^{(2)}$  for each  $i \in [k]$ . We next apply the pigeonhole principle (also known as the bipartite Ramsey theorem) to each of the pairs  $A_x^{(2)}$  and  $A_y^{(2)}$  with  $1 \leq x < y \leq 2k$ ,  $y \neq \pi(x) + k$ , to find the desired subset  $A$ .

To be precise, let  $\{e_1, \dots, e_K\}$  be the set of pairs  $\{x, y\}$  such that  $1 \leq x < y \leq 2k$  and  $y \neq \pi(x) + k$  (recall that  $K = \binom{2k}{2} - k$ ) and, for each  $i \in [2k]$  let  $B_i^{(0)} = A_i^{(2)}$ . We shall define inductively, for each  $i \in [2k]$ , a sequence of sets  $B_i^{(0)} \supset B_i^{(1)} \supset \dots \supset B_i^{(K)}$ . For each pair  $e_\ell = \{x(\ell), y(\ell)\}$  in turn (i.e. for each  $\ell \in [K]$ ), define the sets  $B_i^{(\ell)}$  as follows.

Let  $x' = x'(\ell)$  and  $y' = y'(\ell)$  be the elements matched to  $x = x(\ell)$  and  $y = y(\ell)$  respectively by  $\pi$ , so  $x' = \pi(x) + k$  if  $x \leq k$  and  $x' = \pi^{-1}(x - k)$  if  $x \geq k + 1$ , and similarly for  $y'$ . If  $i \notin \{x, y, x', y'\}$ , then set  $B_i^{(\ell)} = B_i^{(\ell-1)}$ . Let  $B_x^{(\ell)}$  and  $B_y^{(\ell)}$  be the parts of the largest empty bipartite graph induced by  $G$  with  $B_x^{(\ell)} \subset B_x^{(\ell-1)}$ ,  $B_y^{(\ell)} \subset B_y^{(\ell-1)}$  and  $|B_x^{(\ell)}| = |B_y^{(\ell)}|$ . Let  $B_{x'}^{(\ell)}$  be the set  $B_x^{(\ell)} + (k + \pi(i) - i)m$  if  $x \leq k$  and the set  $B_x^{(\ell)} - (k + \pi(i) - i)m$  if  $x \geq k + 1$ . Let  $B_{y'}^{(\ell)}$  be defined from  $B_y^{(\ell)}$  similarly. Note that  $B_{x'}^{(\ell)} \subset B_{x'}^{(\ell-1)}$  and  $B_{y'}^{(\ell)} \subset B_{y'}^{(\ell-1)}$ .

We claim that  $|B_i^{(j)}| \geq S^{(K-j)}(t)$  for each  $i \in [2k]$  and  $j \in [K]$ , and prove it by induction on  $j$ . For  $j = 0$ , the statement is that  $|B_i^{(0)}| \geq S^{(K)}(t) = m'$  for each  $i \in [k]$ , so the base case holds (since each set  $A_i^{(2)}$  has order  $m'$ ). Assume that the result holds for  $j - 1$ . If  $i \notin \{x, y, x', y'\}$ , then  $|B_i^{(j)}| = |B_i^{(j-1)}| \geq S^{(K-j+1)}(t) > S^{(K-j)}(t)$ , so we are done in this case. Now suppose  $i \in \{x, y, x', y'\}$ . By the induction hypothesis,  $|B_\ell^{(j-1)}| \geq S^{(K-j+1)}(t)$  for each  $\ell \in \{x, y\}$  so, by the definition of  $S$ , there exists either a complete bipartite or an empty bipartite graph in  $G[B_x^{(j-1)}, B_y^{(j-1)}]$ , with each part having at least  $S^{(K-j)}(t)$  vertices. Since  $S^{(K-j)}(t) \geq t$ , it cannot be complete (since the sets  $B_x^{(j-1)}$  and  $B_y^{(j-1)}$  are independent), so  $B_\ell^{(j)} \geq S^{(K-j)}(t)$ , for  $\ell \in \{x, y\}$ . Since  $|B_{x'}^{(j)}| = |B_x^{(j)}|$  and  $|B_{y'}^{(j)}| = |B_y^{(j)}|$ , the induction step is complete.

It follows from the claim that  $|B_i^{(K)}| \geq t \geq 1$  for each  $i \in [2k]$ . Observe also that, for each  $i \in [k]$ ,  $B_i^{(K)} + (k + \pi(i) - i)m = B_{\pi(i)+k}^{(K)}$  and that for  $\{i, j\} \in \{e_1, \dots, e_K\}$ , there are no edges in  $G$  between  $B_i^{(K)}$  and  $B_j^{(K)}$ . For each  $i \in [k]$ , choose a vertex  $a(i) \in B_i^{(K)}$ , and let

$a(\pi(i) + k) = a(i) + (k + \pi(i) - i)m \in B_{\pi(i)+k}^{(K)}$ . Let  $A = \{a(1), \dots, a(2k)\}$ . The set  $A$  induces the ordered graph  $H(\pi)$  in  $G$ , so we are done.  $\square$

**Remark 6.** For each permutation  $\pi$ , let  $\tilde{c}(\pi, t)$  denote the smallest constant such that  $|\mathcal{P}_n| < \tilde{c}(\pi, t)^n$  for every  $n \in \mathbb{N}$ , for every hereditary property  $\mathcal{P}$ , satisfying the conditions of the theorem, which avoids  $H(\pi)$ . The bounds given by our proof on the constant  $\tilde{c}(\pi, t)$  are rather large. They could be improved somewhat by choosing the order in which the pairs  $e_i$  are dealt with, and thus obtaining a much stronger inequality than the one we obtained ( $|B_i^{(j)}| \geq S^{(K-j)}(t)$ ), but for simplicity of presentation (and because the actual bounds are not our main interest), we leave this as an exercise for the interested reader. Notice also that, although we assumed that  $K_t \notin \mathcal{P}$ , we only needed  $K_{m'} \notin \mathcal{P}$ .

We finish by noting an immediate consequence of [Theorem 5](#).

**Corollary 15.** *Let  $\mathcal{P}$  be a hereditary property of ordered graphs. If there exists a function  $f : \mathbb{N} \rightarrow \mathbb{N}$  such that*

$$e(G) \leq f(n) = o(n^2)$$

*for every  $G \in \mathcal{P}_n$  and every  $n \in \mathbb{N}$ , then [Conjecture 2](#) holds for  $\mathcal{P}$ .*

**Proof.** Suppose that there is such a function  $f(n) = o(n^2)$ , satisfying  $e(G) \leq f(n)$  for every  $G \in \mathcal{P}_n$  and every  $n \in \mathbb{N}$ . Since  $f(n) = o(n^2)$ , there must exist  $t \in \mathbb{N}$  such that  $e(G) < n^2/4$  for every  $G \in \mathcal{P}_n$  with  $n \geq t$ .

Assume  $t \geq 2$ . Now,  $e(K_t) = \binom{t}{2} \geq \frac{t^2}{4}$  and  $e(K_{t,t}) = t^2$ , so  $K_t \notin \mathcal{P}_t$  and  $K_{t,t} \notin \mathcal{P}_{2t}$ . The result now follows by [Theorem 5](#).  $\square$

## Acknowledgements

The authors would like to thank the anonymous referees for their careful reading of the manuscript and their many helpful comments, which included simplifying the original proof of [Lemma 6](#). They would also like to thank Microsoft Research, where this research was carried out.

The first author was supported during this research by OTKA grant T049398 and US National Science Foundation (NSF) grant DMS-0302804, the second by NSF grant ITR 0225610, and the third by a Van Vleet Memorial Doctoral Fellowship.

## References

- [1] V.E. Alekseev, On the entropy values of hereditary classes of graphs, *Discrete Math. Appl.* 3 (1993) 191–199.
- [2] R. Arratia, On the Stanley–Wilf conjecture for the number of permutations avoiding a given pattern, *Electron. J. Combin.* 6 (1999) 4.
- [3] J. Balogh, B. Bollobás, R. Morris, Hereditary properties of ordered graphs (submitted to a Festschrift in honour of Jaroslav Nešetřil).
- [4] J. Balogh, B. Bollobás, M. Simonovits, On the number of graphs without forbidden subgraphs, *J. Combin. Theory Ser. B* 91 (2004) 1–24.
- [5] J. Balogh, B. Bollobás, D. Weinreich, The speed of hereditary properties of graphs, *J. Combin. Theory Ser. B* 79 (2000) 131–156.
- [6] J. Balogh, B. Bollobás, D. Weinreich, A jump to the Bell number for hereditary graph properties, *J. Combin. Theory Ser. B* 95 (2005) 29–48.

- [7] B. Bollobás, Hereditary properties of graphs: asymptotic enumeration, global structure and colouring, in: *Proceedings of the International Congress of Mathematicians*, vol. III (Berlin, 1998), Doc. Math. 1998, Extra vol. III, 333–342 (electronic).
- [8] B. Bollobás, A. Thomason, Projections of bodies and hereditary properties of hypergraphs, *Bull. London Math. Soc.* 27 (1995) 417–424.
- [9] B. Bollobás, A. Thomason, Hereditary and monotone properties of graphs, in: R.L. Graham, J. Nešetřil (Eds.), *The Mathematics of Paul Erdős*, II, in: *Alg. and Combin.*, vol. 14, Springer-Verlag, New York, Berlin, 1997, pp. 70–78.
- [10] M. Bóna, Exact and asymptotic enumeration of permutations with subsequence conditions, Ph.D. Thesis, MIT (1997).
- [11] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* 2 (1964) 183–190.
- [12] P. Erdős, P. Frankl, V. Rödl, The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs Combin.* 2 (1986) 113–121.
- [13] P. Erdős, D.J. Kleitman, B.L. Rothschild, Asymptotic enumeration of  $K_n$ -free graphs, in: *Colloquio Internazionale sulle Teorie Combinatorie* (Rome, 1973), in: *Atti dei Convegni Lincei* 17, vol. II, Accad. Naz. Lincei, Rome, 1976, pp. 19–27.
- [14] Z. Füredi, P. Hajnal, Davenport–Schinzel theory of matrices, *Discrete Math.* 103 (1992) 233–251.
- [15] C. Hundack, H.J. Prömel, A. Steger, Extremal graph problems for graphs with a color-critical vertex, *Combin. Probab. Comput.* 2 (1993) 465–477.
- [16] M. Klazar, The Füredi–Hajnal conjecture implies the Stanley–Wilf conjecture, in: D. Krob, A.A. Mikhalev, A.V. Mikhalev (Eds.), *Formal Power Series and Algebraic Combinatorics*, Springer, Berlin, 2000, pp. 250–255.
- [17] M. Klazar, Counting pattern-free set partitions I: A generalization of Stirling numbers of the second kind, *European J. Combin.* 21 (2000) 367–378.
- [18] M. Klazar, Counting pattern-free set partitions II: Non-crossing and other hypergraphs, *Electron. J. Combin.* 33 (2000) 737–746.
- [19] M. Klazar, A. Marcus, Extensions of the linear bound in the Füredi–Hajnal conjecture (preprint).
- [20] D.J. Kleitman, K.J. Winston, On the number of graphs without 4-cycles, *Discrete Math.* 41 (1982) 167–172.
- [21] Ph.G. Kolaitis, H.J. Prömel, B.L. Rothschild,  $K_{t+1}$ -free graphs: asymptotic structure and a 0–1 law, *Trans. Amer. Math. Soc.* 303 (1987) 637–671.
- [22] A. Marcus, G. Tardos, Excluded permutation matrices and the Stanley–Wilf conjecture, *J. Combin. Theory Ser. A* 107 (2004) 153–160.
- [23] H.J. Prömel, A. Steger, Excluding induced subgraphs III. A general asymptotic, *Random Structures Algorithms* 3 (1992) 19–31.
- [24] H.J. Prömel, A. Steger, On the asymptotic structure of sparse triangle free graphs, *J. Graph Theory* 21 (1996) 137–151.
- [25] H.J. Prömel, A. Steger, Counting  $H$ -free graphs, *Discrete Math.* 154 (1996) 311–315.
- [26] E.R. Scheinerman, J. Zito, On the size of hereditary classes of graphs, *J. Combin. Theory Ser. B* 61 (1994) 16–39.
- [27] G. Tardos, On 0-1 matrices and small excluded submatrices, *J. Combin. Theory Ser. A* 111 (2005) 266–288.