Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Combinatorial Theory, Series B 95 (2005) 29–48
Journal of Combinatorial Theory Series B
www.elsevier.com/locate/jctb

A jump to the bell number for hereditary graph properties

József Balogh[a], Béla Bollobás[b,c], David Weinreich[d]

[a] *Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA*
[b] *Department of Mathematical Sciences, The University of Memphis, Memphis, TN 38152, USA*
[c] *Trinity College, Cambridge CB2 1TQ, England, UK*
[d] *National Science Foundation, Arlington, VA 22230, USA*

Received 6 September 2001
Available online 31 March 2005

Abstract

A hereditary graph property is a collection of labeled graphs, closed under isomorphism and also under the taking of induced subgraphs. Its speed is the number of graphs in the property as a function of the number of vertices in the graph. Earlier research has characterized the speeds for hereditary graph properties up to $n^{(1+o(1))n}$, and described the properties that have those smaller speeds. The present work provides the minimal speed possible above that range, and gives a structural characterization for properties which exhibit such speeds.

More precisely, this paper sheds light on the jump from below $n^{(1+o(1))n}$ to the range that includes $n^{(1+o(1))n}$. A measure jumps when there are two functions with positive distance such that the measure can take no values between those functions. A clean jump occurs when the bounding functions are well-defined and occur as possible values of the measure. It has been known for some time that the density of a graph jumps; recent work on hereditary graph properties has shown that speeds jump for properties with "large" or "small" speeds.

The current work shows that there is a clean jump for properties with speed in a middle range. In particular, we show that when the speed of a hereditary graph property has speed greater than n^{cn} for

E-mail addresses: jobal@math.ohio-state.edu (J. Balogh), bollobas@msci.memphis.edu (B. Bollobás), dewmath@gmail.com (D. Weinreich).

0095-8956/$ - see front matter
doi:10.1016/j.jctb.2005.02.004

all $c < 1$, the speed is at least $\mathscr{B}_n$, the nth Bell number. Equality occurs only for the property containing all disjoint unions of cliques or its complement.

MSC: 05C

Keywords: Posets; Dilworth's Theorem; Graph properties; Monotone; Hereditary; Speed; Size; Ramsey Theory

1. Introduction

Extremal graph theory concerns itself with the intrinsic structure of graphs, in this sense it is the central field of study in graph theory. Most of the results in the field concern themselves with forcing behavior: that is, the measures studied tend to "jump" in discrete steps. Building on the framework of Turán's Theorem, Erdős and Stone [7] showed that graphs with $t(n, p) + \varepsilon n^2$ edges contain not only the K_{p+1} guaranteed by Turán but $K_{p+1}(t)$, a complete $(p+1)$-partite graph with classes of order t. That is, a graph containing a few (in particular, εn^2) more edges than guarantees a K_{p+1} is forced to contain $K_{p+1}(t)$ as well. This structural result finds a metric counterpoint in the work of Erdős, Stone, and Simonovits (see [7,8]): *Let* $0 < \alpha < 1$, *let* ℓ *be an integer such that* $1 - 1/\ell > \alpha > 1 - 1/(\ell - 1)$, *and let* $m \geqslant 2$ *be an integer. Let* $n = n(m, \alpha - 1/\ell)$ *be sufficiently large. If G has n vertices and* $\alpha\binom{n}{2}$ *edges, then G contains a subgraph on m vertices with at least* $(1 - 1/\ell)\binom{m}{2}$ *edges.*

We shall be concerned with similar discrete steps in a different measure. A *graph property* is an infinite collection of labeled graphs closed under isomorphism. A property is *hereditary* if it is further closed under taking induced subgraphs. The *speed* of a graph property $\mathcal{P}$, denoted $|\mathcal{P}^n|$, is a function of n giving the number of graphs in the property on n vertices. In [4], Bollobás and Thomason showed that the speed of a graph property also jumps, and in precisely the same places as for the number of edges. That is, the only speeds that occur for hereditary graph properties are of the form $|\mathcal{P}^n| = 2^{(1-1/\ell+o(1))\binom{n}{2}}$, for some integer ℓ. More precisely, if a property has speed $2^{(1-1/\ell+\varepsilon+o(1))\binom{n}{2}}$, then it must have speed at least $2^{(1-1/(\ell+1)+o(1))\binom{n}{2}}$. They also presented necessary and sufficient structural characteristics for the properties that evidence each type of speed.

This is the type of result that is ubiquitous in extremal graph theory and often surprising. Formally, let $\mathcal{M}$ be a set of functions. Let f and g be two functions with $\lim \frac{g}{f} = \infty$. We say that $\mathcal{M}$ jumps from f to g if $m \in \mathcal{M}$ and $\limsup \frac{m}{f} = \infty$ implies $m \geqslant g$. We similarly can define a jump from a family of functions $\mathcal{F}$ to a function g. In this case, $\mathcal{F}$ must jump from every $f \in \mathcal{F}$. Note that if $\mathcal{F}$ is defined by asymptotic functions, then $\limsup \frac{m}{f} = \infty$ means that m is greater than f infinitely often. So the Bollobás-Thomason result says that, for any integer ℓ, the set of possible speeds (which we will refer to simply as the speed when our meaning is clear) jumps from $2^{(1-1/\ell+o(1))\binom{n}{2}}$ to $2^{(1-1/(\ell+1)+o(1))\binom{n}{2}}$.

Scheinerman and Zito [16] were the first to note that these jumps also occur with speeds at lower levels. They saw that some classes of functions do not appear as the speed of any hereditary property, and that there are discrete jumps, for example, from polynomial to exponential speeds. As another example, they showed that if a hereditary property has 3 graphs on n vertices for infinitely many values of n, then the speed must in fact jump to a polynomial. In [1], the present authors showed that this jump is in fact to the polynomial $n + 1$. In that paper, we also note many other jumps, enumerate precise functions that form the levels of jumps in two distinct categories, and describe structural characteristics of properties with speeds of each type. Those results are summarized in the following theorem.

Theorem 1. *Let $\mathcal{P}$ be a hereditary property of graphs. Then one of the following is true*:

(1) *there exists $N, k \in \mathbb{N}$ and a collection $\{p_i(n)\}_{i=0}^k$ of polynomials such that for all $n > N$, $|\mathcal{P}^n| = \sum_{i=0}^k p_i(n)i^n$,*
(2) *there exists $k \in \mathbb{N}, k > 1$ such that $|\mathcal{P}^n| = n^{(1-1/k+o(1))n}$,*
(3) $n^{(1+o(1))n} \leqslant |\mathcal{P}^n| \leqslant 2^{o(n^2)}$,
(4) *there exists $k \in \mathbb{N}, k > 1$ such that $|\mathcal{P}^n| = 2^{(1-1/k+o(1))n^2/2}$.*

The existence of jumps within and between the first two cases, proven by the authors in [1], are clear from the statement of the theorem. Jumps within case 4 and the jump to it are, as mentioned above, shown by Bollobás and Thomason [4].

In case 3, however, the behavior is not as clear. Although ay jump from the family of functions in case 2 to some (asymptotically defined) function in case 3 is shown in [3], it is not clear whether there is a lower bound on the functions in case 3. In fact, the behavior of properties in this “penultimate range,” is in general messier than in the rest of the hierarchy. While in all other levels of the hierarchy described in Theorem 1, a possible speed must take on a particular well-defined function, this is not the case in the penultimate range. In [2], the authors show that there exist properties which have speeds that oscillate between extremes. This leads to the question whether any bounds can be given for properties in this range; that is, whether the jump to or from this range is a clean one. This is a reasonable question, as it has been shown in other settings that jumps do not always occur [9].

In this paper, we shall show that the jump from speeds of the type $n^{(1-1/k+o(1))n}$ to the penultimate range is in fact clean, and provide a sharp lower bound, of the Bell numbers, for hereditary properties in this range. In particular, the main result of this paper is the following theorem. It shall be proven in two parts (Theorems 19 and 20) in Section 6. Recall that the nth Bell number, $\mathcal{B}_n$, is the number of partitions of $[n]$ and is asymptotically $\mathcal{B}_n \sim (n/\log n)^n$. $\mathcal{P} = \mathcal{P}_{cl}$ is the property of graphs where each component of each graph in the property is a clique. Its complement $\mathcal{P}_{\overline{cl}}$ consists of Turán graphs (and their induced subgraphs).

Theorem 2. *Let $\mathcal{P}$ be a hereditary graph property. If $|\mathcal{P}^n| \geqslant n^{(1+o(1))n}$, then $|\mathcal{P}^n| \geqslant \mathcal{B}_n$ for all sufficiently large n. Furthermore, equality holds if and only if $\mathcal{P} = \mathcal{P}_{cl}$ or $\mathcal{P} = \mathcal{P}_{\overline{cl}}$.*

The speeds of monotone properties (which are closed under taking arbitrary subgraphs, rather than induced subgraphs) exhibit a similar hierarchy and have a similarly unusual penultimate range [3]. The bound in Theorem 2 will also hold for monotone properties. Analogous results are also possible for hereditary collections of unlabeled graphs, where the lower bound would be $\mathbf{ptn}(n)$, the number of partitions of n. It has been shown that for monotone properties the penultimate range has a clear upper bound [3], and so for that class the bounds on the penultimate range are completely settled. For hereditary properties the upper bound remains unknown.

We shall approach this result by taking a tour through classical graph theory. We start with three important results of combinatorics: Dilworth's Theorem on posets, Ramsey's Theorem on substructures of graphs, and the results of Turán, Erdős, Simonovits, and Stone in extremal graph theory mentioned above. These will be generalized and then applied to hereditary properties of graphs. Definitions and notation will be introduced as needed.

2. Dilworth's Theorem and hypergraphs

To prove the main results of this paper, we shall need some Ramsey-type results on hypergraphs, extending the classical theorem of Dilworth. Other useful Ramsey-type results on graphs, rather than hypergraphs, will be discussed in the next section.

Our definitions are standard, but for completeness shall be given below. A *hypergraph* $\mathcal{H}$ is a pair $\mathcal{H} = (V, \mathbb{E})$, with *vertex set* V and with *edge set* $\mathbb{E}$ consisting of subsets of V. For $x \in V$, the *degree* of x is $d(x) = |\{F : x \in F \in \mathbb{E}\}|$. Clearly, a hypergraph defines a poset on the set of edges, with the order given by inclusion. Viewed this way, we may define the *complement* $\overline{\mathcal{H}}$ of a hypergraph $\mathcal{H} = (V, \mathbb{E})$ by taking the complement of each of the edges over the base set, i.e. $\overline{\mathcal{H}} = (V, \overline{\mathbb{E}})$, where $\overline{\mathbb{E}} = \{V \setminus E : E \in \mathbb{E}\}$. In a hypergraph, we shall allow the empty edge but multiple edges shall not occur.

With this perspective, a *chain* in a hypergraph $\mathcal{H} = (V, \mathbb{E})$ is a collection $\mathbb{F} \subset \mathbb{E}$ such that, for all pairs $A, B \in \mathbb{F}$, either $A \subset B$ or $B \subset A$. An *antichain* in $\mathcal{H}$ is a collection $\mathbb{F} \subset \mathbb{E}$ such that, for all $A, B \in \mathbb{F}$, $A \neq B$, we have $A \not\subset B$. Note that the complement of a chain or an antichain is again a chain or an antichain, respectively.

A basic tool in the theory of posets, Dilworth's Theorem, can be restated for hypergraphs as follows.

Theorem 3. *A hypergraph containing at least $km + 1$ edges contains a chain containing $k + 1$ or an antichain containing $m + 1$ elements.*

While chains have only one form allowed by their definition, antichains are a rich class of sets with very little prescribed form. We wish to extend Dilworth's Theorem to describe the structure of some large antichains that must exist in any hypergraph with no large chain.

Let $\mathcal{H} = (V, \mathbb{E})$ be a hypergraph and $\mathbb{F} = \{A_1, \ldots, A_k\} \subset \mathbb{E}$.

- $\mathbb{F}$ is a *k-star* in $\mathcal{H}$ if there exists $F = \{x_1, \ldots, x_k\} \subset V$ such that $x_i \in A_j$ if and only if $i = j$.
- $\mathbb{F}$ is a *k-costar* in $\mathcal{H}$ if there exists $F = \{x_1, \ldots, x_k\} \subset V$ such that $x_i \notin A_j$ if and only if $i = j$.
- $\mathbb{F}$ is a *k-skewchain* in $\mathcal{H}$ if there exists $F = \{x_1, \ldots, x_{k-1}\} \subset V$ such that $x_i \in A_j$ if and only if $i < j$.

In each case, we call F a *representing set* for $\mathbb{F}$. We may also refer to the pair $(\mathbb{F}, F)$ as the star, costar, or skewchain, where usage should be clear from context. Note that stars and costars are antichains, whereas a skewchain can be a chain, an antichain, or neither. In fact,

$(\mathbb{F}, F)$ is a skewchain if and only if the trace of $\mathbb{F}$ on F is a chain of the same length. Also note that the complement of a star is a costar, while the complement of a skewchain is a skewchain with the same representing set in reversed order.

With these definitions, and motivated by Dilworth's Theorem, we define the number $f(k, \ell, m)$ to be the smallest number such that every hypergraph with at least $f(k, \ell, m)$ edges contains a k-star, ℓ-costar, or m-skewchain. A priori, it is not clear that $f(k, \ell, m)$ is well defined, but the following theorem tells us that this is in fact the case and gives a bound on its growth.

Note that if $\mathcal{H}$ does not contain a k-star, ℓ-costar, or m-skewchain, then $\overline{\mathcal{H}}$ cannot contain an ℓ-star, k-costar, or m-skewchain. Thus $f(k, \ell, m)-1 \geqslant f(\ell, k, m)-1$, and the inequality is also true with k and ℓ reversed. This shows that $f(k, \ell, m)$ is symmetric in k and ℓ.

Theorem 4. *The function $f(k, \ell, m)$ is well defined for all $k, \ell, m \in \mathbb{N}$. In fact, for $k, \ell, m > 2$,*

$$f(k, \ell, m) \leqslant 2(m-1)(m-2) f(k-1, \ell, m) f(k, \ell-1, m) + 1. \tag{1}$$

Proof. Note that $f(k, \ell, m) = 1$ if and only if $\min\{k, \ell, m\} = 1$. When $\min\{k, \ell\} = 2$, the function $f(2, \ell, m) = f(k, 2, m) = m$, as any non-nested pair of edges is both a 2-star and a 2-costar, and $f(k, \ell, 2) = 2$, since repeated edges are not allowed. Hence to prove the theorem it suffices to show Eq. (1) holds when $\min\{k, \ell, m\} > 2$.

Fix $k, \ell, m \geqslant 3$. Let $\mathcal{G} = (V, \mathbb{F})$ be a hypergraph with a number of edges at least as large as the right-hand side of (1). If it contains a chain of length m we are done, since this would also be an m-skewchain. Otherwise, by Theorem 3, $\mathcal{G}$ contains an antichain hypergraph $\mathcal{H} = (V, \mathbb{E})$ with

$$|\mathbb{E}| > 2(m-2) f(k-1, \ell, m) f(k, \ell-1, m).$$

Pick some $x \in V$ that is not in every edge, but is in at least one edge. If $|\mathbb{E}| > d(x) \geqslant |\mathbb{E}|/2$, we shall show that $\mathcal{H}$ contains a k-star, ℓ-costar, or m-skewchain (and, therefore, so does $\mathcal{G}$). If $0 < d(x) < |\mathbb{E}|/2$, then consider the complement of $\mathcal{H}$ and, in this, $|\overline{\mathbb{E}}| > d(x) > |\overline{\mathbb{E}}|/2$. If we can find an ℓ-star, k-costar, or m-skewchain in the complement of $\mathcal{H}$, then $\mathcal{H}$ contains a k-star, ℓ-costar, or m-skewchain. Since $f(k, \ell, m)$ is symmetric in k and ℓ, it does not matter whether we are looking for an ℓ-star or a k-costar or for a k-star or ℓ-costar. Therefore, without loss of generality, we may say $|\mathbb{E}| > d(x) \geqslant |\mathbb{E}|/2$, and showing that $\mathcal{H}$ contains a k-star, ℓ-costar, or m-skewchain will prove the result.

We shall partition the edge set to identify one of the desired structures, establishing a collection of sets according to our choice of x. Let $\mathbb{E}_x = \{E \in \mathbb{E} : x \in E\}$. Since $d(x) \geqslant |\mathbb{E}|/2$,

$$|\mathbb{E}_x| > (m-2) f(k-1, \ell, m) f(k, \ell-1, m).$$

Pick some $A \in \mathbb{E}$ such that $x \notin A$ and let

$$\mathcal{F}_{x,A} = \{E \cap A : E \in \mathbb{E}_x\}.$$

For each $B \in \mathcal{F}_{x,A}$, define

$$\mathbb{E}_B = \{E \in \mathbb{E}_x : E \cap A = B\}.$$

Note that for all $B \in \mathcal{F}_{x,A}$, $A \setminus B \neq \emptyset$, since $\mathbb{E}$ is an antichain.

Case 1: $|\mathbb{E}_B| \geqslant f(k-1, \ell, m)$ *for some* $B \in \mathcal{F}_{x,A}$.

Then $\mathbb{E}_B$ contains either a $(k-1)$-star, an ℓ-costar, or an m-skewchain. In the latter two cases we are done, as $\mathbb{E}_B \subset \mathbb{E}$. Otherwise $\mathbb{E}_B$ contains a $(k-1)$-star, say $\mathbb{S}$, with $\mathbb{S} = \{S_1, \ldots, S_{k-1}\}$ and representing set $S = \{x_1, \ldots, x_{k-1}\}$. As $k \geqslant 3$, and any element in B is in all elements of $\mathbb{E}_B$, we have $x_i \notin B$, so $x_i \notin A$ for all i. However, as noted above, there is some $y \in A \setminus B$. This means that adding A to $\mathbb{S}$ and y to S in the $(k-1)$-star yields a k-star, as desired.

Case 2: $|\mathbb{E}_B| < f(k-1, \ell, m)$ *for all* $B \in \mathcal{F}_{x,A}$.

Then, since the set $\{\mathbb{E}_B : B \in \mathcal{F}_{x,A}\}$ is a partition of $\mathbb{E}_x$,

$$|\mathcal{F}_{x,A}| > (m-2)f(k, \ell-1, m).$$

If $\mathcal{F}_{x,A}$ contains a chain $\mathbb{S}$ with order (at least) $m-1$, then, in a manner similar to that described in Case 1, we may find an m-skewchain in $\mathbb{E}$ consisting of A and a collection of edges each of which intersects A in a different element of $\mathbb{S}$.

Otherwise, by Theorem 3, $\mathcal{F}_{x,A}$ contains an antichain $\mathbb{C}$ with $|\mathbb{C}| > f(k, \ell-1, m)$.

Then, by the induction statement, $\mathbb{C}$ contains either an $(\ell-1)$-costar, a k-star, or an m-skewchain. In the latter two cases we simply take, for each set B in the k-star (respectively, m-skewchain), an edge F from $\mathcal{H}$ with $F \cap A = B$, and the collection we get is a k-star (respectively, m-skewchain) in $\mathcal{H}$, with the same representing set as for $\mathcal{F}_{x,A}$. If $\mathbb{C}$ contains an $(\ell-1)$-costar $\mathbb{S}$ with edges $\{C_1, \ldots, C_{\ell-1}\}$ and representing set $\{x_1, \ldots, x_{\ell-1}\}$, then for each set C_i, there is a $C_i' \in \mathbb{E}_x$ with $C_i = C_i' \cap A$ for all $1 \leqslant i \leqslant \ell-1$. As $C_i' \in \mathbb{E}_x$ for all i, $\{C_1', \ldots, C_{\ell-1}', A\}$ is an ℓ-costar in $\mathcal{H}$ with representing set $\{x_1, \ldots, x_{\ell-1}, x\}$.

Thus, in all cases, a hypergraph with $2(m-1)(m-2)f(k-1, \ell, m)f(k, \ell-1, m)+1$ edges contains a k-star, an ℓ-costar, or an m-skewchain. □

We shall apply this theorem in Section 6 to find certain structures in graphs.

3. Ramsey Theory

Theorem 4 guarantees "large" regular substructures in any hypergraph that is large enough. In this sense, it falls into the vast field of Ramsey Theory. We give the following definitions and notation for clarity and completeness. The advanced reader may skip the following two paragraphs.

Given a graph G, the graph H is isomorphic to an induced subgraph of G if the vertices of H can be mapped to a subset of $V(G)$ so that edges are mapped to edges and non-edges to non-edges. We write $H \leqslant G$, and say H is an *induced subgraph* of G. A vertex set $U \subset V(G)$ induces H if $H \leqslant G$ and the vertices of H can be mapped to U so that edges and non-edges are preserved. We write $H = G[U]$. Slightly differently, given two disjoint sets $U, W \subset V(G)$, the *induced bipartite graph* $G[U, W]$ has vertex set $U \cup W$ and edge set consisting of those edges of G with one end in U and the other end in W. The *bipartite complement* of $G[U, W]$ has the same vertex set as $G[U, W]$ but has edge set $\{uw : u \in U, w \in W, uw \notin E(G)\}$.

We present Ramsey's Theorem here for completeness and notation.

Theorem 5. *There is a number $R(n)$ such that any graph on $R(n)$ vertices contains either K_n or $\overline{K_n}$ as an induced subgraph.*

Ramsey's Theorem says any "large" graph contains an arbitrarily large clique or independent set. We shall be interested in guaranteeing subgraphs other than the complete graph. A Ramsey-type result along these lines for bipartite graphs was obtained by Kővári et al. [11] in response to a question of Zarankiewicz about matrices.

Theorem 6. *Let t be fixed. There is a function $H_t(n) = O(n^{2-1/t})$ such that any bipartite graph on n vertices with at least $H_t(n)$ edges contains $K_{t,t}$ as a subgraph. Further, for $1 \leqslant t < n$,*

$$H_t(n) \leqslant \tfrac{1}{2}(t-1)^{1/t} n^{2-1/t} + \tfrac{1}{2}(t-1)n < 2n^{2-1/t}.$$

Simple calculations give the following corollary.

Corollary 7. *There is a number $n(t)$ such that any bipartite graph with $n(t)$ vertices in each class contains either a $K_{t,t}$ or an independent set containing t vertices from each partition.*

Proof. $n(t) = n$ with $n^2 > 4(2n)^{2-1/t}$ will do. $\square$

Combining Ramsey's Theorem and the result above, we get the following.

Corollary 8. *There is a number $n(t,r)$ such that if G is an $R(r)$-partite graph with $n(t,r)$ vertices in each class then G either contains the Turán graph $T(tr,r)$ or an independent set of tr vertices that intersects r of the sets of the partition in t vertices each.*

Proof. Let G be an $R(r)$-partite graph with $n(n(\ldots(t)\ldots))$ vertices in each class, where the dots signify composition $\binom{R(r)}{2}$ times. Applying Corollary 7 to each pair of partite sets, we obtain an $R(r)$-partite graph with t vertices in each partite set and such that each pair of partite sets is either completely connected or disconnected. The graph H obtained by contracting each partite set to a point is a graph with $R(r)$ vertices and thus contains either a K_r or $\overline{K_r}$, corresponding to $T(tr,r)$ or an independent set that spans r sets of the partition and contains t vertices from each set it intersects, respectively. $\square$

These results concern large substructures that can be guaranteed as a subgraph of an arbitrary graph. We now consider instead a very specific case of both parent and child graphs. Recall that P_n is a path on n vertices. In the next lemma, we show that paths contain highly structured induced path forests (graphs in which every component is a path). The result could be viewed as a statement about colorings of the path, saying that multicolored paths contain induced path forests in which each color appears many times. Or, as stated below, it can be viewed as a statement about words and sentences.

Recall that a *word* is a sequence of letters, where each letter is chosen from a given set, the *alphabet*. A *sentence* can be formed from the word by removing letters and leaving a space wherever consecutive letters have been removed. The *words of the sentence* are then *blocks* of consecutive letters that remain between spaces. This can also be phrased in terms of sequences, colors, subsequences, and blocks, respectively.

For the next result, we fix numbers ℓ, n, and p, and we define $m(\ell, n, p)$ to be the minimal number, if it exists, such that, for any sequence of positive integers $\{a_1, a_2, \ldots, a_p\}$ with $\sum_{i=1}^{p} a_i = n$, every word of length $m(\ell, n, p)$ from an alphabet of size ℓ (containing ℓ letters) contains a sentence with p words, the ith word having length a_i, such that each letter of the alphabet that appears in the sentence appears at least $\lfloor p/\ell \rfloor$ times.

While such a condition sounds quite restrictive, the following lemma says that this number does in fact exist, and gives an inductive bound on its size.

Lemma 9. *The function $m(\ell, n, p)$ is well-defined for all $\ell, n, p \in \mathbb{N}$. Furthermore, for $\ell > 1$,*

$$m(\ell, n, p) \leqslant (p-1)m(\ell-1, n, p) + n + p - 1. \tag{2}$$

Proof. As is implied by the statement, we proceed by induction on ℓ. Clearly $m(1, n, p) = n + p - 1$. So suppose $\ell \geqslant 2$ and $m(\ell-1, n, p)$ exists. Let σ be a word at least as long as given by the right-hand side of (2) built from an alphabet of ℓ letters. Let $\{a_1, a_2, \ldots, a_p\}$ be a set of positive integers with $\sum_{i=1}^{p} a_i = n$. If σ contains a word of length $m(\ell-1, n, p)$ with only $\ell - 1$ different letters, then, by induction, this word contains a sentence with the desired characteristics.

So assume every subword of σ of length $m(\ell-1, n, p)$ contains all ℓ letters. We may construct a sentence with the desired properties with a greedy algorithm. We consider our alphabet to be $[\ell] = \{1, 2, \ldots, n\}$ and, without loss of generality, assume the first letter of σ is 1. Take a word σ_1 of length a_1, and throw out the next entry. If 1 does not appear $\lfloor p/\ell \rfloor$ times in σ_1, skip forward to the next 1 entry. Since every subword of length $m(\ell-1, n, p)$ contains every letter in the alphabet, we need to go forward at most $m(\ell-1, n, p)$ entries. Starting with that 1, take a word σ_2 of length a_2. If the letter 1 has not yet appeared in the sentence $\lfloor p/\ell \rfloor$ times, repeat the process, and repeat for each letter that has not appeared $\lfloor p/\ell \rfloor$ times in the sentence we have picked thus far. We can ensure that our sentence has each letter appearing at least $\lfloor p/\ell \rfloor$ times, as σ has length

$$\begin{aligned}
&(p-1)m(\ell-1, n, p) + n + (p-1) \\
&\quad = (a_1+1) + m(\ell-1, n, p) + (a_2+1) + m(\ell-1, n, p) \\
&\qquad + \ldots + (a_{p-1}+1) + m(\ell-1, n, p) + (a_p). \qquad \square
\end{aligned}$$

4. (ℓ, d)-graphs

We shall introduce more notation. Given $U, W \subset V(G)$, the maximum degree between them, $\Delta(U, W) = \max\{|\Gamma(u) \cap W|, |\Gamma(w) \cap U| : w \in W, u \in U\}$, where $\Gamma(u)$ is the neighborhood of u. With $\overline{\Gamma}(u) = V(G) \setminus (\Gamma(u) \cup \{u\})$, let $\overline{\Delta}(U, W) = \max\{|\overline{\Gamma}(u) \cap W|, |\overline{\Gamma}(w) \cap U| : w \in W, u \in U\}$. Note $\Delta(U, U)$ is simply the maximum degree in $G[U]$, also denoted $\Delta(G[U])$.

A graph H is an *(ℓ, d)-graph* if $V(H)$ admits a partition $V_1, V_2, \ldots, V_\ell$ such that, for each pair i, j (not necessarily distinct) either $\Delta(V_i, V_j) \leqslant d$ or $\overline{\Delta}(V_i, V_j) \leqslant d$. We call $V_1, V_2, \ldots, V_\ell$ an *(ℓ, d)-partition*. It should be clear that, given an (ℓ, d)-partition $V_1, V_2, \ldots, V_\ell$ of H, for each $x \in V(H)$ and $i \in [\ell]$, either $|\Gamma(x) \cap V_i| \leqslant d$ or $|\overline{\Gamma}(x) \cap V_i| \leqslant d$. In the former case we say that x is *sparse* with respect to V_i, in the latter case x is *dense* with respect to V_i. Similarly, if $\Delta(V_i, V_j) \leqslant d$, we say V_i is *sparse* with respect to V_j and if $\overline{\Delta}(V_i, V_j) \leqslant d$, we say V_i is *dense* with respect to V_j. Note that if the sets are large enough (i.e. $\min\{|V_i|\} > 2d$), the terms dense and sparse are mutually exclusive.

We define a *strong* (ℓ, d)-graph to be one which admits an (ℓ, d)-partition, each of whose classes contains at least $5 \cdot 2^\ell d$ vertices. A *strong* (ℓ, d)-partition is defined similarly.

Given any (ℓ', d)-partition, we may obtain another (ℓ, d)-partition (where $\ell' < \ell$) by subdividing any of the classes. Similarly, in some cases we may be able to unify a collection of classes of an (ℓ, d)-partition to obtain a new (ℓ', d)-partition. Hence, (ℓ, d)-partitions of a graph are not unique, even for fixed d. However, if for a fixed d we choose that partition with a minimal number of classes, it may be unique. We would like to find values of d and ℓ that capture the structure of the graph so that the graph admits a unique (ℓ, d)-partition. We shall show that any strong partition can be uniquely modified to yield a unique partition, in a sense that shall be made clear in Theorem 12. First, we need the following two simple lemmas.

Lemma 10. *If two vertices are in the same class of an (ℓ, d)-partition $\{V_i\}_{i=1}^{\ell}$, then the symmetric difference of their neighborhoods has order at most $2\ell d$.*

Proof. Suppose, without loss of generality, $x, y \in V_1$. For each V_i, $|(\Gamma(x) \cap V_i) \cup (\Gamma(y) \cap V_i)| \leqslant 2d$ or $|(\overline{\Gamma}(x) \cap V_i) \cup (\overline{\Gamma}(y) \cap V_i)| \leqslant 2d$, by the definition of (ℓ, d)-partition. Hence, $|\Gamma(x) \Delta \Gamma(y)| \leqslant 2d\ell$. □

If a graph is large, and the value of either d or ℓ is large, then the (ℓ, d)-partition may not reflect the actual patterns of dense/sparse behavior of the graph. Hence, we might consider the following condition on a partition, which says that no two classes of the partition have the same relation to all other classes of the partition. We call this condition $(*)_{\ell,d}$.

If $i, j \in [\ell]$ and $i \neq j$, then there exists a $k \in [\ell]$ such that $\Delta(V_i, V_k) \leqslant d$ and $\overline{\Delta}(V_j, V_k) \leqslant d$ or vice-versa.

Lemma 11. *If two vertices are in different classes of an (ℓ, d)-partition $\{V_i\}_{i=1}^{\ell}$ satisfying $(*)_{\ell,d}$, then the symmetric difference of their neighborhoods has order at least* $\min\{|V_i|\} - 2d$.

Proof. Let $x \in V_1$ and $y \in V_2$. By $(*)_{\ell,d}$, there is a set V_i such that, without loss of generality, x is dense with respect to V_i and y is sparse with respect to V_i. Then, since $|\overline{\Gamma}(x) \cap V_i| \leqslant d$ while $|\Gamma(y) \cap V_i)| \leqslant d$, $|\Gamma(x) \Delta \Gamma(y)| \geqslant |V_i| - 2d$. The result follows. □

Now we may prove our uniqueness result. The theorem provides a "unique" partition for any strong (ℓ, d)-graph H, which we will thereafter call *the unique partition* for H. Although this uniqueness depends on the choice of ℓ and d to some degree, this will not cause difficulties in application.

Theorem 12. *Let H be a strong (ℓ, d)-graph. Then there is an $\ell' \leqslant \ell$ so that H is an (ℓ', pd)-graph which admits a unique (ℓ', pd)-partition π', where $p = \ell - \ell' + 1$. Further, if $k \leqslant \ell'$ and $t \leqslant pd$, then there is no (k, t)-partition of H different from π'.*

Proof. Let H be a strong (ℓ, d)-graph and π an (ℓ, d)-partition. If π does not satisfy $(*)_{\ell,d}$, then there is a pair of classes, without loss of generality, $V_1, V_2 \in \pi$, such that for all $k \in [\ell]$, V_k is dense with respect to both V_1 and V_2 or is sparse with respect to both. In particular, V_1, V_2, and the pair (V_1, V_2) must either be uniformly dense or sparse. But then $V_1 \cup V_2, V_3, \ldots, V_\ell$ is an $(\ell - 1, 2d)$-partition of H, with fewer classes not satisfying $(*)_{\ell-1,2d}$. In this way, given a strong (ℓ, d)-graph H we may join classes that "act the same" in the original partition to obtain an (ℓ', pd)-partition of H, which we will call π', with $p = \ell - \ell' + 1$. Note that π' is strong and satisfies $(*)_{\ell',pd}$.

Suppose $k \leqslant \ell'$ and $t \leqslant pd$, and σ is a (k, t)-partition of H. If σ differs from π', then there are two vertices x and y that are in the same class in σ but in different classes of π'. By Lemma 11, since x and y are in different classes in π', the symmetric difference of their neighborhoods has at least $5 \cdot 2^\ell d - pd$ vertices. However, by Lemma 10, since x and y are in the same class of σ, the symmetric difference of the neighborhoods of x and y is at most $2kt \leqslant 2\ell' pd \leqslant 3\ell pd - pd < 5 \cdot 2^\ell d - pd$, a contradiction. Hence $\sigma = \pi'$ or there is no (k, t)-partition.

By the proof above, a strong graph admits a unique partition with a minimal number of sets. Hence we will call this *the minimal partition* of H. While the minimality (of the number of sets) and the uniqueness of the partition depend on the initial choice of ℓ and d, this does not cause any complications in the applications below. While we do not mention ℓ or d in our usage of these terms, it should be understood that the partition is only minimal/unique for the choice of ℓ and d.

Given a strong (ℓ, d)-graph H, consider its minimal partition $(V_1, \ldots, V_{\ell'})$. We define $\varphi(H)$ as the graph we obtain from the minimal partition of H by replacing $H[V_i, V_j]$ with its bipartite complement for every pair if $\Delta(V_i, V_j) > \ell d$. The maximal degree of $\varphi(H)$ is at most ℓd (although the minimal partition may be an (ℓ', pd)-partition, we can achieve the same $\varphi(H)$ by switching analogous classes in the (ℓ, d)-partition, and thus the result holds). The function $\varphi(H)$ depends on d and ℓ, but not on the minimal partition, since the minimal partition is uniquely determined for a strong (ℓ, d)-graph.

As a corollary to Theorem 12, we can see that strong (ℓ, d)-graphs not only produce a unique minimal partition, but that the unique partition is preserved by any subgraph that would still be strong. Recalling that the minimal partition satisfies $(*)$ and is strong, the proof is analogous to that of Theorem 12. We therefore omit it.

Corollary 13. *Let H be a strong (ℓ, d)-graph with unique partition $V_1, \ldots, V_{\ell'}$. Let $F \leqslant \varphi(H)$ with $|V(F) \cap V_i| \geqslant 5 \cdot 2^\ell d$ for all i such that $V(F) \cap V_i \neq \emptyset$. Let $G = H[V(F)]$. Then G is a strong (ℓ, d)-graph with unique partition $V(F) \cap V_1, \ldots, V(F) \cap V_{\ell'}$. Further, $\varphi(G)$ is well-defined and is equal to F.*

This brings us to our most important result for the counting of graphs and determining of speeds of hereditary properties.

Lemma 14. *If H is a strong (ℓ, d)-graph, then $Aut(H) \subseteq Aut(\varphi(H))$. Therefore, the number of distinct labelings of H is at least as large as the number of distinct labelings of $\varphi(H)$.*

Proof. We shall show that any automorphism of H is also an automorphism of $\varphi(H)$. Let Ψ be an automorphism of H and $\pi = (V_i, \dots, V_{\ell'})$ be the unique partition of H. First we claim Ψ preserves the classes of π (up to the labeling). This follows from Lemmas 10 and 11; i.e. two vertices are in the same class if and only if the cardinality of the symmetric difference of their neighborhoods is at most $2\ell d$ (which is clearly $< 5 \cdot 2^{\ell} d$). Hence, without loss of generality, we can suppose that for all i and all $x \in V_i$, we have $\Psi(x) \in V_i$. Let $x \in V_i$, $y \in V_j$, where the case $i = j$ is included. Then $\Psi(x) \in V_i$ and $\Psi(y) \in V_j$. Without loss of generality assume $xy \in E(H)$. Then $\Psi(x)\Psi(y) \in E(H)$ as well.

If V_i is sparse with respect to V_j, then $xy \in E(\varphi(H))$, and we have to prove that $\Psi(x)\Psi(y) \in E(\varphi(H))$. Indeed, as $\Psi(x)\Psi(y) \in E(H)$ and V_i is sparse to V_j, the map φ keeps $\Psi(x)\Psi(y) \in E(\varphi(H))$. Similarly, if V_i is dense with respect to V_j then $xy \notin E(\varphi(H))$, and we have to prove that $\Psi(x)\Psi(y) \notin E(\varphi(H))$. As $xy \in E(H)$, we have $\Psi(x)\Psi(y) \in E(H)$, and because (V_i, V_j) is a dense pair, $\Psi(x)\Psi(y) \notin E(\varphi(H))$ follows. □

Note that the converse of the statement in the proof above is generally not true. Two vertices that have different neighborhoods in H may have identical neighborhoods in $\varphi(H)$.

Also note that if H is not strong, the lemma may not be true at all. In particular, if a graph is an (ℓ, d)-graph, it is also an $(\ell+1, d)$-graph, with an $(\ell+1, d)$-partition obtained by breaking up any class of the (ℓ, d)-partition. As a trivial example, consider K_{20} as a $(2, 1)$-graph. Consider the non-trivial $(2, 1)$-partition into V_1 and V_2, with $|V_1| = 3$. Then $2 = \Delta(V_1, V_1) \leqslant 2, 16 = \Delta(V_2, V_2) > 2$ and $17 = \Delta(V_1, V_2) > 2$, so $\varphi(K_{20}) = K_3 \dot\cup \overline{K_{17}}$, which has $\binom{20}{3}$ different labelings while K_{20} has only one. In fact, for this reason, $\varphi(H)$ is only defined for strong (ℓ, d)-graphs (many smaller examples with $|\mathrm{Aut}(\varphi(G))| \neq |\mathrm{Aut}(G)|$ would not be strong). The lemma applies, however, when we view K_{20} as a strong $(2, 1)$-graph. Then with the unique (and trivial) partition, $|\mathrm{Aut}(\varphi(K_{20}))| = |\mathrm{Aut}(K_{20})| = 1$.

5. Hereditary properties of graphs

The terminology of (ℓ, d)-graphs may seem a bit awkward, but in fact (ℓ, d)-graphs are critical to understanding the structure of complicated properties of graphs.

In [1], the present authors show that partitions like those in an (ℓ, d)-graph provide an easy way to bound the number of graphs in a property. In fact, in a certain range, where $|\mathcal{P}^n|$ is roughly factorial in n, they provide the best way to bound the speed. These results will be presented below, but as usual we need a few more definitions.

Given a graph G and collection of vertices $v_1, \dots, v_t \in V(G)$, we say that the disjoint sets $U_1, \dots, U_m \subset V(G)$ are *distinguished by* $X = \{v_1, \dots, v_t\}$ if, for each i, every vertex of U_i has the same neighborhood in X and for each $i \neq j$, $x \in U_i$, $y \in U_j$ implies x and y have different neighborhoods in X. We say X *distinguishes* U_i. The set X is a *minimal* distinguishing set if no proper subset of X distinguishes the same sets.

The following definition is new, and may seem odd at first. Let $k_{\mathcal{P}}$ be the minimum k, if it exists, such that there is an $m > 0$ for which no $G \in \mathcal{P}$ contains a set of vertices that distinguish m sets, each of order at least k. If no such k exists, set $k_{\mathcal{P}} = \infty$. This definition allows us to connect hereditary properties to (ℓ, d)-graphs in the following surprising result [1, Lemma 27].

Lemma 15. *If $\mathcal{P}$ is a hereditary property with $k_{\mathcal{P}} < \infty$, then there exist absolute constants $\ell_{\mathcal{P}}$ and $c_{\mathcal{P}}$ such that for all $G \in \mathcal{P}$, the graph G contains an induced subgraph H such that H is an $(\ell_{\mathcal{P}}, k_{\mathcal{P}})$-graph and $|V(G \setminus H)| < c_{\mathcal{P}}$.*

More importantly for computing speeds, we also showed the following [1, Theorem 28].

Theorem 16. *Let $\mathcal{P}$ be a hereditary property with $k_{\mathcal{P}} < \infty$. Then $|\mathcal{P}^n| \geqslant n^{(1+o(1))n}$ if and only if for all m there exists a strong $(\ell_{\mathcal{P}}, k_{\mathcal{P}})$-graph H in $\mathcal{P}$ such that $\varphi(H)$ has a component of order at least m.*

We shall put these ideas to use in the next section, where we deal with properties at the bottom of the penultimate range.

First, an easy pair of technical lemmas. Recall that P_n is a path on n vertices.

Lemma 17. *Let $D > 2$. If G is connected and $\Delta(G) \leqslant D$, then, for any $n \leqslant \log_D |V(G)|$, we have $P_n \leqslant G$.*

Proof. Pick any $v \in V(G)$. The number of vertices at distance d from v is at most D^d, by the degree condition. Hence, for any $n \geqslant 1$, the number of vertices at distance less than n from v is at most $\sum_{i=0}^{n-1} D^i < D^n$. If $n \leqslant \log_D |V(G)|$, then $|V(G)| > D^n$ so there must be a $u \in V(G)$ with $d(u, v) \geqslant n$. An n vertex subpath of a shortest u-v path is then an induced P_n in G. $\square$

It is not too surprising that there is a relationship between the speed of a property and the structure of graphs in that property. We can count the number of labelings of a graph roughly by grouping vertices into classes that can be distinguished from each other and then choosing labels for a group en masse. For example, if a graph G consists only of disjoint cliques, then the number of labelings of G is $\binom{n}{c_1, c_2, \ldots, c_m}$, where $c_1, \ldots, c_m$ are the orders of the cliques. Continuing the argument, consider the property, $\mathcal{P}_{cl}$, where each component of each graph in the property is a clique. Then a labeled graph in $\mathcal{P}_{cl}$ on n vertices can be described by an unordered partition of $[n]$; in fact, there is a one-to-one correspondence between graphs of $\mathcal{P}_{cl}^n$ and such partitions of $[n]$. Hence $|\mathcal{P}_{cl}^n| = \mathcal{B}_n$.

In fact, this property, $\mathcal{P}_{cl}$, and its complement $\mathcal{P}_{\overline{cl}}$ consisting of Turán graphs (and their induced subgraphs), are the only properties with speed exactly $\mathcal{B}_n$, as shall be shown in Section 6.

If the groupings of vertices are not cliques, then clearly such a count gives only a lower bound. Such lower bounds are instructive, however, in examining the speeds that occur.

We now consider a collection of graphs that is not a hereditary property, but will appear as a subcollection of many of the graph properties we shall consider in Section 6.

A *path forest* is a graph in which every component is a path. Let $p_i(n)$ be the number of labeled path forests on n vertices that have i components. As we noted above, $\mathcal{B}_n$ is the number of labeled graphs in which every component is a clique. It is clear that the number of labeled graphs with i components such that each component is a clique must be less than the number of such graphs with path components, and thus we have $\sum_{i=1}^{n} p_i(n) > \mathcal{B}_n$. In fact, the Bell number is dominated by any term in the sum that corresponds to path forests with fewer than $\sqrt{n}$ parts, as shall be shown in the following lemma. This is not too surprising, as $p_1(n)$ is the number of cyclic permutations of n, which clearly dominates the number of partitions of n.

Lemma 18. *If $n > c^2$, then $p_c(n) > \mathcal{B}_n$.*

Proof. How many path forests are there on $[n]$ with c parts? In order to form a labeled path forest, we could start with any permutation of $[n]$ and break it into c parts by splitting it at $c-1$ places. This does not necessarily yield a unique path forest: any rearrangement of the paths would allow a different permutation of $[n]$ to give the same labeled path forest, as would reversing the direction of any non-trivial path. Hence each path forest with c parts can be represented by at most $c!2^c$ different permutations of $[n]$.

As mentioned in the discussion before the proof, $p_1(n) = n!/2 > \mathcal{B}_n$. Also, $p_2(n) = n!(n-1)/8 > B_n$ since $n > c^2$. Finally, if $c \geqslant 3$,

$$p_c(n) \geqslant \frac{n!\binom{n-1}{c-1}}{2^c c!} = n!\left(\frac{n-1}{(c-1)2c}\right)\left(\frac{n-2}{(c-2)2(c-1)}\right)\cdots\left(\frac{n-c+1}{2\cdot 2}\right)\frac{1}{2}$$
$$\geqslant n!\left(\frac{n-1}{(c-1)2c}\right)^{c-2}\frac{n-c+1}{8} \geqslant n!\frac{c^2-c+2}{8} \geqslant n! > \mathcal{B}_n,$$

proving the assertion. □

6. A lower bound on the penultimate range

We are now ready to prove our main results. We shall prove Theorem 2 in two parts, considering separately properties where $k_{\mathcal{P}} < \infty$ and where $k_{\mathcal{P}} = \infty$. The hard work of the first case has been done by the collection of lemmas and theorems in the preceding sections.

Theorem 19. *Let $\mathcal{P}$ be a hereditary property with $|\mathcal{P}^n| \geqslant n^{(1+o(1))n}$. If $k_{\mathcal{P}} < \infty$, then, for n sufficiently large, $|\mathcal{P}^n| > \mathcal{B}_n$.*

Proof. Let $\ell_{\mathcal{P}}$ be given by Lemma 15. Let $c = (2\ell_{\mathcal{P}}^2 k_{\mathcal{P}} + 2k_{\mathcal{P}}\ell_{\mathcal{P}} + 1)\ell_{\mathcal{P}}$ and assume $n > c^2$. By Theorem 16, for all m, $\mathcal{P}$ contains a strong $(\ell_{\mathcal{P}}, k_{\mathcal{P}})$-graph H such that $\varphi(H)$ has a component of order m. Since $\varphi(H)$ has bounded degree and m is arbitrarily large, Lemma 17 says $\mathcal{P}$ contains a graph H such that $\varphi(H)$ contains a path of length $m(\ell_{\mathcal{P}}, n, c)$, where m is the function from Lemma 9. Color the vertices of this path according to the minimal $(\ell_{\mathcal{P}}, k_{\mathcal{P}})$-partition of H. According to Lemma 9, this path contains any path-forest of total length n and c components in a way that each class of the partition of H is intersected at least $c/\ell_{\mathcal{P}} = 2\ell_{\mathcal{P}}^2 k_{\mathcal{P}} + 2k_{\mathcal{P}}\ell_{\mathcal{P}} + 1$ times. For each of these path forests F, Corollary 13 guarantees a graph $G_F \leqslant H$ such that $\varphi(G_F) = F$. Since $\mathcal{P}$ is hereditary, $G_F \in \mathcal{P}$ for all such path forests F. Now let $\mathcal{F}$ be the collection of all labeled path forests with n vertices and c components. With our choice of n, Lemma 18 says $|\mathcal{F}| > \mathcal{B}_n$, and, by Lemma 14, $|\mathcal{P}^n| > |\mathcal{F}|$, since each graph in $\mathcal{F}$ is the image of some graph in $\mathcal{P}^n$ under φ. $\square$

The proof for when $k_{\mathcal{P}} = \infty$ involves case analysis of the structures that might occur and an application of the results of Section 3. Note that, both by a theorem of [1] and independently by the theorem below, $k_{\mathcal{P}} = \infty$ implies $|\mathcal{P}^n| \geqslant n^{(1+o(1))n}$.

Theorem 20. *Let $\mathcal{P}$ be a hereditary property. If $k_{\mathcal{P}} = \infty$ then $|\mathcal{P}^n| \geqslant \mathcal{B}_n$. Equality holds if and only if $\mathcal{P} = \mathcal{P}_{cl}$ or $\mathcal{P}_{\overline{cl}}$.*

Proof. Fix n. We shall show that $|\mathcal{P}^n| \geqslant \mathcal{B}_n$ and note the restrictive criteria for equality. Let k and r be large enough to guarantee that the Ramsey results we apply below hold, and let $m \geqslant f(r, r, r)$, where $f(r, r, r)$ is the function from Theorem 4.

By the definition of $k_{\mathcal{P}}$, for all k, m, there is a $G \in \mathcal{P}$ and $X \subseteq V(G)$ such that X distinguishes m sets each of order at least k. Let $G \in \mathcal{P}$ be such a graph for our choices of k and m. Let X be a distinguishing set for G and $V_1, \ldots, V_m$ be distinguished sets of order at least k. Let $\mathcal{H} = (X, \mathbb{E})$ be the hypergraph defined by $\mathbb{E} = \{\Gamma_X(V_i)\}$. That is, the vertices of $\mathcal{H}$ are the distinguishing vertices of G and the edges correspond to the subsets of X that create the distinguished partition. Note that $\mathcal{H}$ has no multiple edges, so $|\mathbb{E}| = m$. Hence, by our choice of m, $\mathcal{H}$ contains an r-star, r-costar, or an r-skewchain. Consider the induced subgraph $S' \leqslant G$ corresponding to this r-star, r-costar, or r-skewchain.

The graph S' contains a set, X, of r vertices which distinguish r sets, $V_1, \ldots, V_r$, each of order at least k. As k and r were chosen large enough, Theorem 5 of Ramsey guarantees that each of $X, V_1, \ldots, V_r$ contains either a large clique or a large independent set. Similarly, Corollary 8 guarantees that among the distinguished sets, ignoring their internal structure, there is a large spanning independent set or a large Turán graph. Thus we may first choose a distinguishing set $S_1 \subseteq X$ so that S_1 is either K_n or $\overline{K_n}$ and let (with perhaps appropriate renumbering) $V_1, \ldots, V_n$ be those sets distinguished by S_1. For each V_i, let $U_i \subseteq V_i$ such that each U_i is either K_n or $\overline{K_n}$ uniformly and so that all pairs (U_i, U_j) induce either $K_{n,n}$ or $\overline{K_{n,n}}$ uniformly.

Let $S \subseteq S'$ be the graph induced by S_1 and $\bigcup_{i=1}^n U_i$. Its distinguishing set is S_1 and let $S_2 = V(S) \setminus S_1 = \bigcup_{i=1}^n U_i$. Then $|S_1| = n$ and $|S_2| = n^2$. Further, S_1 is either K_n or $\overline{K_n}$ and S_2 is one of K_{n^2}, $\overline{K_{n^2}}$, nK_n, or $\overline{nK_n} = T(n^2, n)$.

Based on the manner in which S is created as a subgraph of S', we know that the hypergraph based on the distinguishing relationship between S_1 and S_2 is either an n-star, n-costar, or n-skewchain. There are 24 different possible structures that can be described as above, and $\mathcal{P}$ must contain an arbitrarily large graph containing one of these structures. The 8 possibilities when the hypergraph based on S is an n-star are shown in Fig. 1; there are similarly 8 possible structures if that hypergraph is a costar, and 8 more when it is a skewchain.

If $S_2 = nK_n$ or $\overline{nK_n}$, then $\mathcal{P}_{cl}^n \subseteq \mathcal{P}^n$ or $\mathcal{P}_{\overline{cl}}^n \subseteq \mathcal{P}^n$, respectively, and $|\mathcal{P}^n| \geqslant \mathcal{B}_n$, since $|\mathcal{P}_{cl}^n| = \mathcal{B}_n$ as noted earlier. Note that, for these cases, equality occurs if and only if the

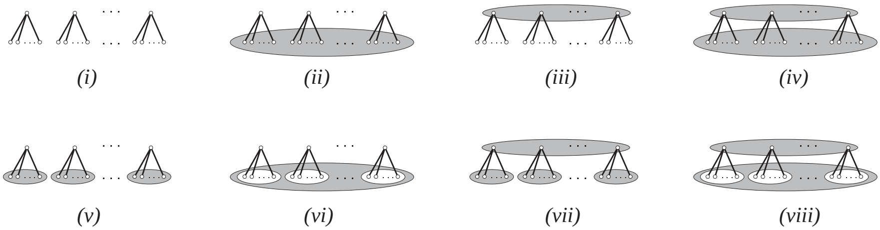

Fig. 1. The eight possibilities for S if the hypergraph based on $G \in \mathcal{P}$ contains an r-star. The gray ovals indicate sets which induce a clique, while an empty oval within a grey oval represents an induced independent set within an otherwise fully connected group of vertices (i.e. a Turán graph). In each figure, the top vertices are $S_1 = \{v_1, \ldots, v_n\}$ and the bottom vertices are $S_2 = U_1 \cup U_2 \cup \cdots \cup U_n$.

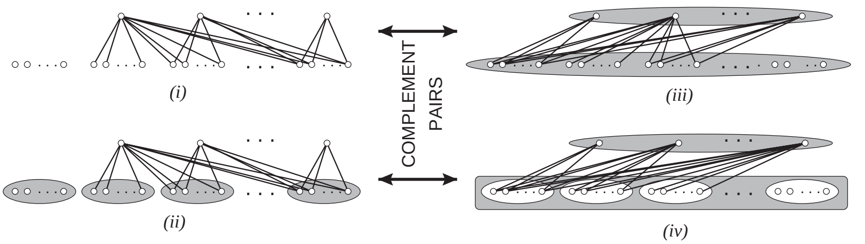

Fig. 2. The four possibilities for S if a hypergraph based on $G \in \mathcal{P}$ contains an r-skewchain, but the property does not contain $\mathcal{P}_{cl}$ or its complement. The horizontal pairings indicate complementary pairs of graphs. The gray ovals indicate sets which induce a clique, while an empty oval within a gray oval represents an induced independent set within an otherwise fully connected group of vertices (i.e. a Turán graph). In each figure, the top vertices are $S_1 = \{v_1, \ldots, v_{n-1}\}$ and the bottom vertices are $S_2 = U_0 \cup U_1 \cup U_2 \cup \cdots \cup U_{n-1}$.

property is $\mathcal{P}_{cl}$ or $\mathcal{P}_{\overline{cl}}$. We shall show that in the other cases (when $\mathcal{P}_{cl} \not\subseteq \mathcal{P}$) the inequality is always strict.

We need only consider the cases when $S_2 = K_{n^2}$ or $\overline{K_{n^2}}$. For the configurations based on stars, these are the top 4 structures shown in Fig. 1. Considering costars and skewchains, then, there are 12 possibilities in total to be considered. We may cut this number in half by counting the number of labelings of the complementary property $\overline{\mathcal{P}} = \{\overline{G} : G \in \mathcal{P}\}$. Clearly $|\overline{\mathcal{P}}^n| = |\mathcal{P}^n|$. Each possible configuration based on a star is the complement of a configuration based on a costar, and the 4 remaining configurations based on a skewchain may be paired as shown in Fig. 2, so we need only consider the left partner of each pair in that figure. The 2 skewchains and 4 chains give us 6 cases to consider.

In each of the cases, we shall show that $|\mathcal{P}^n| > \mathcal{B}_n$ by finding a correspondence between subgraphs of S on n vertices (which, since $\mathcal{P}$ is hereditary, are graphs in $\mathcal{P}$) and partitions of $[n]$.

We describe a function from partitions of $[n]$ to subgraphs of S as follows, which we will refer to as f. Call a class of a partition *non-trivial* if it has at least 2 elements; a *singleton* is a vertex forming a trivial class.

Let π be a partition of $[n]$ and let $\pi = A_1, A_2, \ldots, A_t, B$, where each A_i is non-trivial, B is the collection of singletons in π, and the A_i are ordered so that $\min A_i < \min A_j$

whenever $i < j$. Let $a_i = \min A_i$ and $B_i = A_i \setminus \{a_i\}$ for all *i*. We define a function, g, as follows. Let $g(a_i) = v_i$ and $g(B_i) \subseteq U_i$. Since *S* is unlabeled and the vertices of U_i are indistinguishable, this is well-defined. If *S* is like Fig. 1(iv), let $g(B) \subseteq \bigcup_{i>t} U_i$, otherwise let $g(B) \subseteq \overline{K_n} \leq \bigcup_{i>r}(\{v_i\} \cup U_i)$. Let *f* be the function that maps π to the subgraph of *S* induced by the image of *g*, with each vertex labeled by its preimage under *g*. Since *B* was mapped to a set of vertices that are indistinguishable in $f(\pi)$, the choice of vertices for $g(B)$ (and the choice of $\overline{K_n}$) does not matter.

Our strategy will be to show that *f* is an injection from partitions of $[n]$ into graphs of $\mathcal{P}^n$. Given a graph that is the image of a partition of $[n]$ under *f*, we shall uniquely reconstruct that partition. In some cases, however, we shall need to modify *f* before applying this strategy, in others we shall enumerate the few exceptions that cannot be reconstructed and count them separately. These sub-strategies shall be made clear in the cases below.

Case 1: *S is as in Fig.* 1(i). In this case, the image of any partition is a star forest, and we can reconstruct the partition according to its components. As the smallest label in a set of the partition always gets mapped to the center of the star, no two partitions give the same star forest, so *f* is a injection. Further, some labeled star forests (and hence subgraphs of *S*) are not the images of any partition (e.g. the star $K_{1,n-1}$ with center labeled *n*), so the inequality is strict.

Case 2: *S is as in Fig.* 1(ii). To reconstruct the partition, we need to consider only non-isolated vertices, as isolated vertices must correspond to singletons in the original partition. For each non-isolated vertex, we consider the number of maximal cliques that it is a member of, where by *maximal clique* we mean a clique that is not a proper subset of any clique. If each of these vertices is in only one maximal clique, then this graph is the image of a partition with only one non-trivial set. Otherwise, all vertices that are members of only one maximal clique are the smallest elements of their set in the partition, and the clique partition of the graph that they induce is the partition that yielded the graph. Once again we may find subgraphs of *S* that are not the images of any partition. Indeed, any labeled subgraph of *S* with more than one maximal clique that gives the label *n* to some non-isolate that is in only one maximal clique (some element of that clique must be in more than one maximal clique) is not the image of any partition. Therefore, the inequality is strict.

Case 3: *S is as in Fig.* 1(iii). To reconstruct the partition, we do the same as in Case 2 in reverse. If every vertex is in only one maximal clique, then this graph is the image of a partition with at most one non-trivial class, this class having order 2. Otherwise, the non-trivial component of the graph has at least 4 vertices, and the vertices with degree greater than 1 induce a clique. This clique is S_1, and induces (by neighborhoods) a partition of the graph into stars. This, in turn, corresponds to the original partition. Again, many subgraphs of *S* are missed (e.g. for any $k > 1$, any labeling of the graph consisting of a *k*-clique with a pendant edge and $n - k - 1$ isolated vertices), so the inequality is strict.

Case 4: *S is as in Fig.* 1(iv). Recall that in this case we use a slightly different definition of *f* to account for the lack of isolates in *S*. Here the isolates get mapped to vertices in S_2 after the labels in the non-trivial parts are mapped. The function *f*, defined either way, is not an injection. However, it is "almost" an injection; we shall isolate those configurations which do not have a unique preimage under *f* and show that enough subgraphs of *S* are not in the image of *f* to account for the overlap.

We again proceed by identifying the maximal cliques in the graph. Note that no vertex of S is in more than two maximal cliques, so in any induced subgraph no vertex is in more than two maximal cliques either. Also note that the function f will always yield a connected graph.

There are no isolated vertices in S, so we cannot immediately identify the trivial sets of the partition. If every vertex is in only one maximal clique, the graph must be a complete graph, since it is a subgraph of S. The graph might be the image of either the discrete or indiscrete partition. This is one of three cases where two partitions get mapped to the same graph. Each of these will be identified below and all will be dealt with at the conclusion of this case.

If the graph has more than one maximal clique, then the partition that yielded the graph must have at least one non-trivial part that is not all of $[n]$. In the case that the partition has exactly one non-trivial part, the graph will have exactly two maximal cliques. If the partition has at least two non-trivial parts, then the graph will have at least 4 maximal cliques.

Suppose the graph has exactly two maximal cliques. Consider the set, M, of all vertices appearing in both cliques. There must be a vertex u such that one of the two cliques is $M \cup \{u\}$, where the label on u is smaller than any label in M and $M \cup \{u\}$ corresponds to the non-trivial part of the partition that yielded the graph. If one of the cliques has more than one element outside of M, or has an element outside of M with a label larger than one appearing in M, then u, and hence the partition, is uniquely determined. Hence, the only way that $M \cup \{u\}$ is not uniquely determined is if $M = \{3, \ldots, n\}$. Then the partition that yielded the graph is either $\{\{1, 3, 4, \ldots, n\}\{2\}\}$ or $\{\{2, \ldots, n\}\{1\}\}$. This is the second case of two partitions yielding the same graph, and shall again be dealt with below.

As noted above, if the graph has more than two maximal cliques, it has at least 4, and the partition that yielded it has at least two non-trivial parts. Any vertex corresponding to an element in a non-trivial part appears in two maximal cliques: S_1 or S_2 and the clique corresponding to its part in the partition. Hence if a graph has at least 4 maximal cliques and some vertex appears in only one maximal clique, then it corresponds to an isolate in the original partition, the maximal clique it is a member of is S_2, and the partition may be uniquely reconstructed according to the cliques that intersect S_2.

So let us assume that the graph under consideration has at least 4 maximal cliques and each vertex is in two maximal cliques. Thus the partition that yielded it has no singleton sets and is not the indiscrete partition.

If the graph contains at least 5 maximal cliques, then, since the graph is a subgraph of S, there are 2 non-intersecting cliques which partition the vertices (each of the other cliques intersect both of these two cliques, but not each other). These non-intersecting cliques are S_1 and S_2, and the partition may be reconstructed corresponding to the other maximal cliques.

Thus (except for the two cases deferred above) the partitions and the graphs they map to are in 1–1 correspondence, unless the graph has exactly 4 maximal cliques where every vertex appears in exactly two maximal cliques. Consider such a graph H. Since H has more than two maximal cliques, any partition that yields H must have at least two non-trivial parts, and since there are fewer than 5 maximal cliques such a partition must have no more than two non-trivial parts. Hence a partition that yields H must be a two-set partition of $[n]$ with no singletons (the latter condition as no vertex is in only one maximal clique). Therefore H has a clique with only two vertices, corresponding to S_1. If only one clique

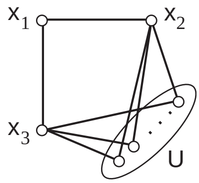

Fig. 3. A troublesome configuration for an image under *f*.

has exactly two vertices, then these vertices correspond to the smallest elements of their respective parts in the partition, and we may reconstruct the partition according to the other cliques they are in. Thus the only case left to consider is that shown in Fig. 3, where two different maximal cliques each have exactly two vertices.

Assume here $n > 4$. Then $|U| > 1$ so the vertices of U must have come from S_2 and x_1 must then have come from S_1. Hence, the label on x_1 must be smaller than the label on either x_2 or x_3. If x_1 is only smaller than one of them, then the partition is known. For example, if label$(x_2) <$ label(x_1) but label$(x_3) >$ label(x_1), then $\{\{x_1, x_3\}, \{x_2\} \cup U\}$ is the original partition. So assume that the label on x_1 is smaller than the labels on both x_2 and x_3. Then label$(x_1)= 1$, since all the labels on U must be bigger than one of label(x_2) or label(x_3). Similarly, the label 2 can only occur on x_2 or x_3, so without loss of generality the label on x_2 is 2. Now if the label on x_3 is not 3, then some vertex in U is 3 and the label on x_3 is bigger than 3, so again $\{\{x_1, x_3\}, \{x_2\} \cup U\}$ must be the original partition. If the label on x_3 is in fact 3, then we cannot be sure whether the graph is the image of $\{\{x_1, x_3\}, \{x_2\} \cup U\}$ or $\{\{x_1, x_2\}, \{x_3\} \cup U\}$. These pairs and the two pairs mentioned earlier get mapped to the same graph.

We shall modify *f* so that every partition is mapped to a unique graph in $\mathcal{P}$ in an invertible fashion.

Map the partition $\{\{1, 2\}, \{3, \ldots, n\}\}$ to its image under *f*, but map $\{\{1, 3\}, \{2, 4, 5, \ldots, n\}\}$, which under *f* gets mapped to the same image, to the graph in Fig. 3 so x_1 is labeled 3, x_2 is labeled 1, and x_3 is labeled 2. This latter graph is not the image under *f* of any partition, as the label on x_1 is not smaller than that on x_2 or x_3. Similarly, we map the indiscrete partition as usual but the discrete partition to the graph in Fig. 3 so that x_1 is labeled n, x_2 is labeled $n - 1$, and x_3 is labeled $n - 2$. Map $\{\{1, 3, 4, \ldots, n\}\{2\}\}$ to its normal image under *f* but map $\{\{2, \ldots, n\}\{1\}\}$ to the same graph (an $n-2$ clique joined to two independent vertices) with the two external vertices labeled n and $n - 1$. Once again, this latter graph is not the image of any partition, as may be seen by the arguments in the paragraph discussing that case. This new function is clearly invertible.

Again, we have missed many graphs, including any graph isomorphic to that in Fig. 3 where $\{x_1, x_2, x_3\}$ is labeled from any set other than $\{1, 2, 3\}$ or $\{n - 2, n - 1, n\}$ (among others). Hence the inequality is strict.

Case 5: *S is as in Fig.* 2(i). We shall refer to the labeling described in the caption. To reconstruct the partition, we consider simply the degrees of the vertices. All vertices of degree 0 are singletons in the original partition. If a vertex has degree 1, it is either in U_1 or is v_p, where p is the number of non-trivial parts in the partition (that is, the last of the vertices in S_1). In the latter case, the last non-singleton set in the partition must have exactly two elements.

We may determine which vertices belong to U_1 as follows. If any pair of vertices with degree 1 have a common neighbor, then they are both in U_1 and their common neighbor is v_1, and its neighbors of degree 1 constitute U_1. We may completely reconstruct the partition by considering the neighbors of v_1 of degree 2. These constitute U_2 and their common neighbor other than v_1 is v_2. We may continue in this fashion to reconstruct all sets U_i and thus determine the original partition.

If no pair of vertices with degree 1 has a common neighbor, then there can be at most two vertices of degree 1. If there is only one of degree 1, this must be the entirety of U_1, its neighbor is v_1, and we may proceed as above. Otherwise there are exactly two vertices with degree 1 (and the first and last non-trivial set in the partition each have exactly two elements). Call these vertices x and y, and their neighbors x' and y', respectively. Then either the label on x is bigger than the label on x' and x' is v_1, or the label on y is bigger than the label on y', and y' is v_1. Once we have identified v_1, we may proceed as above to reconstruct the partition. Note that the possibilities for x and y above are exclusive. We may construct a class of labeled graphs that are not the images of any partition, and thus obtain a strict inequality, by considering the last case and, for example, labeling two vertices of degree 1 with labels 1 and 2.

Case 6: *S is as in Fig.* 2(ii). We proceed as in the previous case, but, for non-isolates, rather than consider the degree of each vertex we consider the number of maximal cliques it is in. With this change, the argument is identical, and the same type of example described there shows that the inequality is strict. □

Taken together, these results give Theorem 2 as a corollary, as promised in the introduction. Thus, the penultimate range has a clear and sharp lower bound, and properties jump to this lower bound from case 2 of Theorem 1. We have settled a major mystery about a difficult region of the speed hierarchy, but there is still much to discover about this range.

7. Structure of minimal penultimate properties

In the past, we have sought to give, in addition to bounds on the speed of properties, collections of minimal properties that "force" the speed to be in the range given. For the penultimate range, this type of result is only partially done.

Let G_1 be the infinite graph with structure given in Fig. 1(i), i.e. an infinite forest of infinite stars. Similarly, define $G_2, G_3, \dots, G_6$ as the infinite graphs corresponding to Figs. 1(ii)–1(iv), 2(i), and 2(ii), respectively. Let $\mathcal{P}(G_i)$ be the property containing all finite induced subgraphs of G_i. Then the proof of Theorem 20 implies that under the hypotheses of that theorem, $\mathcal{P}$ contains one of $\{\mathcal{P}_{cl}, \mathcal{P}(G_1), \mathcal{P}(G_2), \dots, \mathcal{P}(G_6)\}$ or its complement. However, these are not the minimal properties for the penultimate range, as it is only when $k_{\mathcal{P}} = \infty$ that we can guarantee the inclusion. It would seem that a characterization of minimal properties for $k_{\mathcal{P}} < \infty$ would not have a simple representation, although surely there is such a class. This class of minimal properties would have to be based on φ-transformations of path forests, and we would be happy to see such a result in the future.

The space provided by the strict inequality in Theorem 19, which is due to Lemma 18, does tell us that the very smallest of properties in this range, however, do in fact contain one of the properties listed above. In particular, if $\mathcal{B}_n \leqslant |P^n| < \sqrt{n}\mathcal{B}_n$, then $\mathcal{P}$ must contain one of these properties, the upper bound given by the bound on $p_c(n)$. This itself may be a jump, and further study is warranted.

In fact, it is unclear whether there are jumps within the penultimate range at any point between its bounds. This promises to be a rich area of research in the future.

Acknowledgments

The authors would like to thank the anonymous referees who provided several useful comments on the presentation of this work.

References

[1] J. Balogh, B. Bollobás, D. Weinreich, The speed of hereditary properties of graphs, J. Combin. Theory Ser. B 79 (2) (2000) 131–156.
[2] J. Balogh, B. Bollobás, D. Weinreich, The penultimate range of growth for graph properties, European J. Combin. 22 (3) (2001) 277–289.
[3] J. Balogh, B. Bollobás, D. Weinreich, Measures on monotone properties of graphs, Discrete Appl. Math. 116 (1–2) (2002) 17–36.
[4] B. Bollobás, A. Thomason, Hereditary and monotone properties of graphs, The mathematics of Paul Erdős, II, Algorithms and Combinatorics, vol. 14, Springer, Berlin, 1997, pp. 70–78.
[7] P. Erdős, A.H. Stone, On the structure of linear graphs, Bull. Amer. Math. Soc. 52 (1946) 1087–1091.
[8] P. Erdős, M. Simonovits, A limit theorem in graph theory, Studia Sci. Mat. Hung. Acad. 1 (1966) 51–57.
[9] P. Frankl, V. Rodl, Hypergraphs do not jump, Combinatorica 4 (2–3) (1984) 149–159.
[11] T. Kővári, V.T. Sós, P. Turán, On a problem of K. Zarankiewicz, Colloq. Math. 3 (1954) 50–57.
[16] E.R. Scheinerman, J. Zito, On the size of hereditary classes of graphs, J. Combin. Theory (B) (1994) 16–39.