

DECIDING THE BELL NUMBER FOR HEREDITARY GRAPH PROPERTIES*

AISTIS ATMINAS[†], ANDREW COLLINS[‡], JAN FONIOK[§], AND VADIM V. LOZIN[‡]

Abstract. The paper [J. Balogh, B. Bollobás, D. Weinreich, *J. Combin. Theory Ser. B*, 95 (2005), pp. 29–48] identifies a jump in the speed of hereditary graph properties to the Bell number B_n and provides a partial characterization of the family of minimal classes whose speed is at least B_n . In the present paper, we give a complete characterization of this family. Since this family is infinite, the decidability of the problem of determining if the speed of a hereditary property is above or below the Bell number is questionable. We answer this question positively by showing that there exists an algorithm which, given a finite set \mathcal{F} of graphs, decides whether the speed of the class of graphs containing no induced subgraphs from the set \mathcal{F} is above or below the Bell number. For properties defined by infinitely many minimal forbidden induced subgraphs, the speed is known to be above the Bell number.

Key words. hereditary class of graphs, speed of hereditary properties, Bell number, decidability

AMS subject classifications. 05C30, 05C75, 68C05

DOI. 10.1137/15M1024214

1. Introduction. A *graph property* (or a *class of graphs*¹) is a set of graphs closed under isomorphism. Given a property \mathcal{X} , we write \mathcal{X}_n for the number of graphs in \mathcal{X} with vertex set $\{1, 2, \dots, n\}$ (that is, we are counting *labeled* graphs). Following [5], we call \mathcal{X}_n the *speed* of the property \mathcal{X} .

A property is *hereditary* if it is closed under taking induced subgraphs. It is well known (and can be easily seen) that a graph property \mathcal{X} is hereditary if and only if \mathcal{X} can be described in terms of forbidden induced subgraphs. More formally, for a set \mathcal{F} of graphs we write $\text{Free}(\mathcal{F})$ for the class of graphs containing no induced subgraph isomorphic to any graph in the set \mathcal{F} . A property \mathcal{X} is hereditary if and only if $\mathcal{X} = \text{Free}(\mathcal{F})$ for some set \mathcal{F} . We call \mathcal{F} a set of *forbidden induced subgraphs* for \mathcal{X} and say that graphs in \mathcal{X} are *\mathcal{F} -free*.

The speeds of hereditary properties and their asymptotic structure have been extensively studied, originally in the special case of a single forbidden subgraph [9, 10, 12, 14, 15, 16], and more recently in general [1, 4, 5, 6, 7, 17]. These studies showed, in particular, that there is a certain relationship between the speed of a property \mathcal{X} and the structure of graphs in \mathcal{X} , and that the rates of the speed growth constitute discrete layers. The first four lower layers have been distinguished in [17]:

*Received by the editors June 3, 2015; accepted for publication (in revised form) March 24, 2016; published electronically May 17, 2016. The authors gratefully acknowledge support from DIMAP: the Centre for Discrete Mathematics and its Applications at the University of Warwick, and from EPSRC, grant EP/L020408/1. A conference version of this paper appeared in the Proceedings of the 40th International Workshop on Graph-Theoretic Concepts in Computer Science (WG 2014), Lecture Notes Comput. Sci. 8747.

<http://www.siam.org/journals/sidma/30-2/M102421.html>

[†]School of Science and Technology, Nottingham Trent University, Clifton Campus, Nottingham NG11 8NS, UK (aistis.atminas@ntu.ac.uk).

[‡]DIMAP and Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK (A.Collins.2@warwick.ac.uk, V.Lozin@warwick.ac.uk).

[§]School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Manchester M1 5GD, UK (j.foniok@mmu.ac.uk).

¹Throughout the paper we use the two terms—graph property and class of graphs—interchangeably.

these are constant, polynomial, exponential, and factorial layers. In other words, the authors of [17] showed 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.

Independently, similar results were obtained by Alekseev in [2]. Moreover, Alekseev provided the first four layers with the description of all minimal classes, that is, he identified in each layer the family of all classes every proper hereditary subclass of which belongs to a lower layer (see also [5] for some more involved results). In each of the first four lower layers the set of minimal classes is finite and each of them is defined by finitely many forbidden induced subgraphs. This provides an efficient way of determining whether a property \mathcal{X} belongs to one of the first three layers.

One more jump in the speed of hereditary properties was identified in [7] and it separates—within the factorial layer—the properties with speeds strictly below the Bell number B_n from those whose speed is at least B_n . With a slight abuse of terminology we will refer to these two families of graph properties as properties below and above the Bell number, respectively. The importance of this jump is due to the fact that all the properties below the Bell number are well-structured. In particular, all of them have bounded clique-width [3] and all of them are well-quasi-ordered by the induced subgraph relation [13]. From the results in [5, 13] it follows that every hereditary property below the Bell number can be characterized by finitely many forbidden induced subgraphs and hence the membership problem for each of them can be decided in polynomial time.

Even so, very little is known about the boundary separating the two families, that is, very little is known about the *minimal* classes above the Bell number. Paper [7] distinguishes two cases in the study of this question: the case where a certain parameter associated with each class of graphs is finite and the case where this parameter is infinite. In the present paper, we call this parameter the *distinguishing number*. For the case where the distinguishing number is infinite, [7] provides a complete description of minimal classes, of which there are precisely 13. For the case where the distinguishing number is finite, [7] mentions only one minimal class above the Bell number (linear forests) and leaves the question of characterizing other minimal classes open.

In the present paper, we give a complete answer to the above open question: we provide a structural characterization of all minimal classes above the Bell number with a finite distinguishing number. This family of minimal classes is infinite, which makes the problem of deciding whether a hereditary class is above or below the Bell number questionable. Nevertheless, for properties defined by *finitely many* forbidden induced subgraphs, our characterization allows us to prove decidability of this problem: we show that there exists an algorithm which, given a finite set \mathcal{F} of graphs, decides whether the class $\text{Free}(\mathcal{F})$ is above or below the Bell number.

All preliminary information related to the topic of the paper can be found in section 2. In section 3, we describe the minimal classes above the Bell number. Finally, in section 4 we present our decidability result. Section 5 concludes the paper with an open problem.

2. Preliminaries and preparatory results.

2.1. Basic notation and terminology. All graphs we consider are undirected without multiple edges. The graphs in our hereditary classes have no loops; however, we allow loops in some auxiliary graphs, called “density graphs” and denoted usually by H , that are used to represent the global structure of our hereditary classes.

If G is a graph, $V(G)$ stands for its vertex set, $E(G)$ for its edge set, and $|G|$ for the number of vertices (the *order*) of G . The edge joining two vertices u and v is uv (we do not use any brackets); uv is the same edge as vu .

If $W \subseteq V(G)$, then $G[W]$ is the subgraph of G induced by W . For W_1, W_2 disjoint subsets of $V(G)$ we define $G[W_1, W_2]$ to be the bipartite subgraph of G with vertex set $W_1 \cup W_2$ and edge set $\{uv : u \in W_1, v \in W_2, uv \in E(G)\}$. The *bipartite complement* of $G[W_1, W_2]$ is the bipartite graph in which two vertices $u \in W_1, v \in W_2$ are adjacent if and only if they are not adjacent in $G[W_1, W_2]$.

The *neighborhood* $N(u)$ of a vertex u in G is the set of all vertices adjacent to u , and the *degree* of u is the number of its neighbors. Note that if (and only if) there is a loop at u then $u \in N(u)$.

As usual, $P_n, C_n,$ and K_n denote the path, the cycle, and the complete graph with n vertices, respectively. Furthermore, $K_{1,n}$ is a star (i.e., a tree with $n + 1$ vertices one of which has degree n), and $G_1 + G_2$ is the disjoint union of two graphs. In particular, mK_n is the disjoint union of m copies of K_n .

A *forest* is a graph without cycles, i.e., a graph every connected component of which is a tree. A *star forest* is a forest every connected component of which is a star, and a *linear forest* is a forest every connected component of which is a path.

A *quasi-order* is a binary relation which is reflexive and transitive. A *well-quasi-order* is a quasi-order which contains neither infinite strictly decreasing sequences nor infinite antichains (sets of pairwise incomparable elements). That is, in a well-quasi-order any infinite sequence of elements contains an infinite increasing subsequence.

Recall that the Bell number B_n , defined as the number of ways to partition a set of n labeled elements, satisfies the asymptotic formula $\ln B_n/n = \ln n - \ln \ln n + \Theta(1)$.

Balogh, Bollobás, and Weinreich [7] showed that if the speed of a hereditary graph property is at least $n^{(1-o(1))n}$, then it is actually at least B_n ; hence we call any such property a *property above the Bell number*. Note that this includes hereditary properties whose speed is exactly equal to the Bell numbers (such as the class of disjoint unions of cliques).

2.2. (ℓ, d) -graphs and sparsification. Given a graph G and two vertex subsets $U, W \subseteq V(G)$, define $\Delta(U, W) = \max\{|N(u) \cap W|, |N(w) \cap U| : u \in U, w \in W\}$. With $\bar{N}(u) = V(G) \setminus (N(u) \cup \{u\})$, let

$$\bar{\Delta}(U, W) = \max\{|\bar{N}(u) \cap W|, |\bar{N}(w) \cap U| : w \in W, u \in U\}.$$

Note that $\Delta(U, U)$ is simply the maximum degree in $G[U]$.

DEFINITION 2.1. *Let G be a graph. A partition $\pi = \{V_1, V_2, \dots, V_{\ell'}\}$ of $V(G)$ is an (ℓ, d) -partition if $\ell' \leq \ell$ and for each pair of not necessarily distinct integers $i, j \in \{1, 2, \dots, \ell'\}$ either $\Delta(V_i, V_j) \leq d$ or $\bar{\Delta}(V_i, V_j) \leq d$. We call the sets V_i bags. A graph G is an (ℓ, d) -graph if it admits an (ℓ, d) -partition.*

If $\Delta(V_i, V_j) \leq d$, we say V_i is d -sparse with respect to V_j , and if $\bar{\Delta}(V_i, V_j) \leq d$, we say V_i is d -dense with respect to V_j . We will also say that the pair (V_i, V_j) is d -sparse or d -dense, respectively. Note that if the bags are large enough (i.e., $\min\{|V_i|\} > 2d + 1$), the terms d -dense and d -sparse are mutually exclusive.

DEFINITION 2.2. *A strong (ℓ, d) -partition is an (ℓ, d) -partition each bag of which contains at least $5 \times 2^\ell d$ vertices; a strong (ℓ, d) -graph is a graph which admits a strong (ℓ, d) -partition.*

Given any strong (ℓ, d) -partition $\pi = \{V_1, V_2, \dots, V_{\ell'}\}$ we define an equivalence relation \sim on the bags by putting $V_i \sim V_j$ if and only if for each k , either V_k is

d -dense with respect to both V_i and V_j , or V_k is d -sparse with respect to both V_i and V_j . Let us call a partition π *prime* if all its \sim -equivalence classes are of size 1. If the partition π is not prime, let $p(\pi)$ be the partition consisting of unions of bags in the \sim -equivalence classes for π .

We proceed to showing that the partition $p(\pi)$ of a strong (ℓ, d) -graph does not depend on the choice of a strong (ℓ, d) -partition π . The following three lemmas are the ingredients for the proof of this result.

LEMMA 2.3. *Consider any strong (ℓ, d) -graph G with any strong (ℓ, d) -partition π . Then $p(\pi)$ is an $(\ell, \ell d)$ -partition with at least $5 \times 2^\ell d$ vertices in each bag.*

Proof. Consider two bags $W_1, W_2 \in p(\pi)$. By definition $W_i = \bigcup_{s \in S_i} V_s$ for some $S_i \subset \{1, 2, \dots, \ell'\}$, $i = 1, 2$. Also, by the definition of the partition, for all $(s_1, s_2) \in S_1 \times S_2$ the pairs (V_{s_1}, V_{s_2}) are all either d -dense or d -sparse. If they are d -sparse, then for any $s_1 \in S_1$ we have $\Delta(V_{s_1}, W_2) \leq \sum_{s_2 \in S_2} \Delta(V_{s_1}, V_{s_2}) \leq |S_2|d$. Since this holds for every $s_1 \in S_1$, for all $x \in W_1$ we have that $|N(x) \cap W_2| \leq |S_2|d$. Similarly we conclude that for all $x \in W_2$ we have $|N(x) \cap W_1| \leq |S_1|d$. Therefore, $\Delta(W_1, W_2) \leq \max(|S_1|, |S_2|)d \leq \ell d$. If the pairs of bags are d -dense, a similar argument proves that $\overline{\Delta}(W_1, W_2) \leq \ell d$. Hence the partition $p(\pi)$ is an $(\ell, \ell d)$ -partition. As it is obtained by unifying some bags from a strong (ℓ, d) -partition, we conclude that each bag is of size at least $5 \times 2^\ell d$. \square

LEMMA 2.4 (see [7, Lemma 10]). *Let G be a graph with an (ℓ, d) -partition π . If two vertices $x, y \in G$ are in the same bag V_k , then the symmetric difference of their neighborhoods $N(x) \ominus N(y)$ is of size at most $2\ell d$.*

LEMMA 2.5. *Let G be a graph with a strong (ℓ, d) -partition π . If two vertices $x, y \in V(G)$ belong to different bags of the partition $p(\pi)$, then the symmetric difference of their neighborhoods $N(x) \ominus N(y)$ is of size at least $5 \times 2^\ell d - 2d$.*

Proof. Take any two vertices $x \in V_i$ and $y \in V_j$ with bags V_i and V_j belonging to different \sim -equivalence classes. Then there is a bag V_k such that one of the pairs (V_i, V_k) and (V_j, V_k) is d -dense and the other one is d -sparse; without loss of generality, suppose (V_i, V_k) is d -sparse and (V_j, V_k) is d -dense. Then, in particular, $|N(x) \cap V_k| \leq d$ and $|N(y) \cap V_k| \geq |V_k| - d$. Hence $|N(x) \ominus N(y)| \geq |N(y) \setminus N(x)| \geq |V_k| - 2d \geq 5 \times 2^\ell d - 2d$. \square

We are now ready to prove the uniqueness of $p(\pi)$.

THEOREM 2.6. *Let G be a strong (ℓ, d) -graph with strong (ℓ, d) -partitions π and π' . Then $p(\pi) = p(\pi')$.*

Proof. Assume two vertices $x, y \in V(G)$ are in the same bag of the partition $p(\pi)$. By Lemma 2.3, $p(\pi)$ is an $(\ell, \ell d)$ -partition, so applying Lemma 2.4 to $p(\pi)$ we obtain $|N(x) \ominus N(y)| \leq 2\ell(\ell d) = 2\ell^2 d < 5 \times 2^\ell d - 2d$. Thus by Lemma 2.5, x and y are in the same bag of $p(\pi')$. Hence, using symmetry, x and y are in the same bag of $p(\pi)$ if and only if they are in the same bag of $p(\pi')$. We deduce that the partitions are the same, i.e., $p(\pi) = p(\pi')$. \square

With any strong (ℓ, d) -partition $\pi = \{V_1, V_2, \dots, V_{\ell'}\}$ of a graph G we can associate a *density graph* (with loops allowed) $H = H(G, \pi)$: the vertex set of H is $\{1, 2, \dots, \ell'\}$ and there is an edge joining i and j if and only if (V_i, V_j) is a d -dense pair (so there is a loop at i if and only if V_i is d -dense).

For a graph G , a vertex partition $\pi = \{V_1, V_2, \dots, V_{\ell'}\}$ of G , and a graph H (with loops allowed) with vertex set $\{1, 2, \dots, \ell'\}$, we define (as in [5]) the H, π -transform $\psi(G, \pi, H)$ to be the graph obtained from G by replacing $G[V_i, V_j]$ with its bipartite

complement for every pair (V_i, V_j) for which ij is an edge of H , and replacing $G[V_i]$ with its complement for every V_i for which there is a loop at the vertex i in H .

Moreover, if π is a strong (ℓ, d) -partition we define $\phi(G, \pi) = \psi(G, \pi, H(G, \pi))$. Note that π is a strong (ℓ, d) -partition for $\phi(G, \pi)$ and each pair (V_i, V_j) is d -sparse in $\phi(G, \pi)$. We now show that the result of this “sparsification” does not depend on the initial strong (ℓ, d) -partition.

PROPOSITION 2.7. *Let G be a strong (ℓ, d) -graph. Then for any two strong (ℓ, d) -partitions π and π' , the graph $\phi(G, \pi)$ is identical to $\phi(G, \pi')$.*

Proof. Suppose that $\pi = \{U_1, U_2, \dots, U_{\ell}\}$ and $\pi' = \{V_1, V_2, \dots, V_{\ell'}\}$. By Theorem 2.6, $p(\pi) = p(\pi') = \{W_1, W_2, \dots, W_{\ell''}\}$. Consider two vertices x, y of G . Let i, j, i', j', i'', j'' be the indices such that $x \in U_i, x \in V_{i'}, x \in W_{i''}, y \in U_j, y \in V_{j'}, y \in W_{j''}$. As the partitions have at least $5 \times 2^{\ell}d$ vertices in each bag, ℓd -dense and ℓd -sparse are mutually exclusive properties. Hence the pair (U_i, U_j) is d -sparse if and only if $(W_{i''}, W_{j''})$ is ℓd -sparse if and only if $(V_{i'}, V_{j'})$ is d -sparse, and analogously for dense pairs. Therefore $ij \in E(H(G, \pi))$ if and only if $i''j'' \in E(H(G, p(\pi)))$ if and only if $i'j' \in E(H(G, \pi'))$. We conclude that xy is an edge of $\phi(G, \pi)$ if and only if it is an edge of $\phi(G, \pi')$. \square

Proposition 2.7 motivates the following definition, originating from [5].

DEFINITION 2.8. *For a strong (ℓ, d) -graph G , its sparsification is $\phi(G) = \phi(G, \pi)$ for any strong (ℓ, d) -partition π of G .*

2.3. Distinguishing number $k_{\mathcal{X}}$. In this section, we discuss the distinguishing number of a hereditary graph property, which is an important parameter introduced by Balogh, Bollobás, and Weinreich in [5].

Given a graph G and a set $X = \{v_1, \dots, v_t\} \subseteq V(G)$, we say that the disjoint subsets U_1, \dots, U_m of $V(G)$ are *distinguished* by X if for each i , all vertices of U_i have the same neighborhood in X , and for each $i \neq j$, vertices $x \in U_i$ and $y \in U_j$ have different neighborhoods in X . We also say that X *distinguishes* the sets U_1, U_2, \dots, U_m .

DEFINITION 2.9. *Given a hereditary property \mathcal{X} , we define the distinguishing number $k_{\mathcal{X}}$ as follows:*

- (a) *If for all $k, m \in \mathbb{N}$ we can find a graph $G \in \mathcal{X}$ that admits some $X \subset V(G)$ distinguishing at least m sets, each of size at least k , then put $k_{\mathcal{X}} = \infty$.*
- (b) *Otherwise, there must exist a pair (k, m) such that any vertex subset of any graph $G \in \mathcal{X}$ distinguishes at most m sets of size at least k . We define $k_{\mathcal{X}}$ to be the minimum value of k in all such pairs.*

In [5], Balogh, Bollobás, and Weinreich show that the speed of any hereditary property \mathcal{X} with $k_{\mathcal{X}} = \infty$ is above the Bell number. To study the classes with $k_{\mathcal{X}} < \infty$, in the next sections we will need two results from their paper.

LEMMA 2.10 (see [5, Lemma 27]). *If \mathcal{X} is a hereditary property with finite distinguishing number $k_{\mathcal{X}}$, then there exist absolute constants $\ell_{\mathcal{X}}, d_{\mathcal{X}} \leq k_{\mathcal{X}}$ and $c_{\mathcal{X}}$ such that for all $G \in \mathcal{X}$, the graph G contains an induced subgraph G' such that G' is an $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph and $|V(G) \setminus V(G')| < c_{\mathcal{X}}$.*

By removing all the small bags with fewer than $5 \times 2^{\ell_{\mathcal{X}}}d_{\mathcal{X}}$ vertices, which affects only the constant $c_{\mathcal{X}}$, we can actually assume that the graph G' is a *strong* $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph. This observation allows us to strengthen Lemma 2.10 as follows.

LEMMA 2.11. *If \mathcal{X} is a hereditary property with finite distinguishing number $k_{\mathcal{X}}$, then there exist absolute constants $\ell_{\mathcal{X}}$, $d_{\mathcal{X}}$, and $c_{\mathcal{X}}$ such that for all $G \in \mathcal{X}$, the graph G contains an induced subgraph G' such that G' is a strong $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph and $|V(G) \setminus V(G')| < c_{\mathcal{X}}$.*

Finally, we will use the following theorem.

THEOREM 2.12 (see [5, Theorem 28]). *Let \mathcal{X} be a hereditary property with $k_{\mathcal{X}} < \infty$. Then $\mathcal{X}_n \geq n^{(1+o(1))n}$ if and only if for every m there exists a strong $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph G in \mathcal{X} such that its sparsification $\phi(G)$ has a connected component of order at least m .*

3. Structure of minimal classes above Bell. In this section, we describe minimal classes with speed above the Bell number. In [7], Balogh, Bollobás, and Weinreich characterized all minimal classes with infinite distinguishing number. In section 3.1 we report this result and prove additionally that each of these classes can be characterized by finitely many forbidden induced subgraphs. Then in section 3.2 we move on to the case of finite distinguishing number, which had been left open in [7].

3.1. Infinite distinguishing number.

THEOREM 3.1 (Balogh–Bollobás–Weinreich [7]). *Let \mathcal{X} be a hereditary graph property with $k_{\mathcal{X}} = \infty$. Then \mathcal{X} contains at least one of the following (minimal) classes:*

- (a) *the class \mathcal{K}_1 of all graphs each of whose connected components is a clique;*
- (b) *the class \mathcal{K}_2 of all star forests;*
- (c) *the class \mathcal{K}_3 of all graphs whose vertex set can be split into an independent set I and a clique Q so that every vertex in Q has at most one neighbor in I ;*
- (d) *the class \mathcal{K}_4 of all graphs whose vertex set can be split into an independent set I and a clique Q so that every vertex in I has at most one neighbor in Q ;*
- (e) *the class \mathcal{K}_5 of all graphs whose vertex set can be split into two cliques Q_1, Q_2 so that every vertex in Q_2 has at most one neighbor in Q_1 ;*
- (f) *the class \mathcal{K}_6 of all graphs whose vertex set can be split into two independent sets I_1, I_2 so that the neighborhoods of the vertices in I_1 are linearly ordered by inclusion (that is, the class of all chain graphs);*
- (g) *the class \mathcal{K}_7 of all graphs whose vertex set can be split into an independent set I and a clique Q so that the neighborhoods of the vertices in I are linearly ordered by inclusion (that is, the class of all threshold graphs);*
- (h) *the class $\overline{\mathcal{K}}_i$ of all graphs whose complement belongs to \mathcal{K}_i as above, for some $i \in \{1, 2, \dots, 6\}$ (note that the complementary class of \mathcal{K}_7 is \mathcal{K}_7 itself).*

As an aside, it is perhaps worth noting that each of the minimal classes admits an infinite universal graph. To be specific, \mathcal{K}_1 is the age (the class of all finite induced subgraphs) of U_1 , the disjoint union of ω cliques, each of order ω . The remaining universal graphs are depicted in Figure 1; a grey oval indicates a clique (of order ω).

Aiming to prove that each of the classes above is defined by forbidding finitely many induced subgraphs, we first state an older result by Földes and Hammer about split graphs which we make use of in our proof. A *split graph* is a graph whose vertex set can be split into an independent set and a clique.

THEOREM 3.2 (see [11]). *The class of all split graphs is exactly the class $\text{Free}(2K_2, C_4, C_5)$.*

Before showing the characterization of the classes \mathcal{K}_1 – \mathcal{K}_6 in terms of forbidden

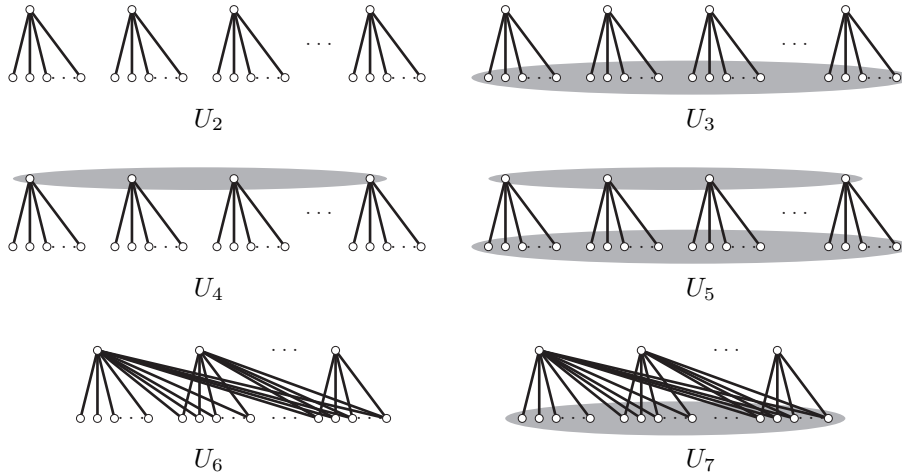


FIG. 1. The universal graphs.

induced subgraphs, we introduce some of the less commonly appearing graphs: the claw $K_{1,3}$, the 3-fan F_3 , the diamond K_4^- , and the graph H_6 (Figure 2).

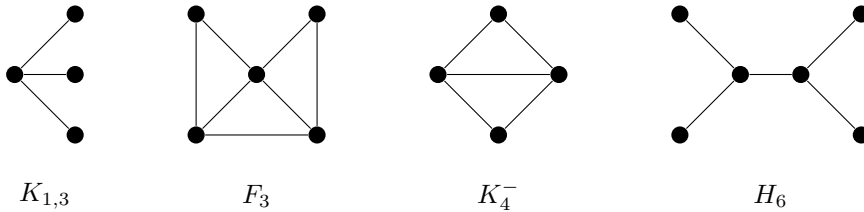


FIG. 2. Some small graphs.

THEOREM 3.3. *Each of the classes of Theorem 3.1 is defined by finitely many forbidden induced subgraphs.*

Proof. First, observe that if we define $\bar{\mathcal{X}}$ as the class of the complements of all graphs in \mathcal{X} , then $\text{Free}(\mathcal{F}) = \text{Free}(\bar{\mathcal{F}})$. Hence if each class \mathcal{K}_i is defined by finitely many forbidden induced subgraphs, then so is each $\bar{\mathcal{K}}_i$.

(a) $\mathcal{K}_1 = \text{Free}(P_3)$: It is trivial to check that P_3 does not belong to \mathcal{K}_1 , and any graph not containing an induced P_3 must be a collection of cliques.

(b) $\mathcal{K}_2 = \text{Free}(K_3, P_4, C_4)$: Obviously, none of the graphs K_3, P_4, C_4 belongs to \mathcal{K}_2 . Let $G \in \text{Free}(K_3, P_4, C_4)$. Since every cycle of length at least 5 contains P_4 , G does not contain any cycles; thus G is a forest. The absence of a P_4 implies that the diameter of any connected component of G is at most 2, hence, G is a star forest.

(c) $\mathcal{K}_3 = \text{Free}(\mathcal{F})$ for $\mathcal{F} = \{2K_2, C_4, C_5, K_{1,3}, F_3\}$: It is easy to check that none of the forbidden graphs belong to \mathcal{K}_3 . Let $G \in \text{Free}(\mathcal{F})$. By Theorem 3.2, G is a split graph. Split G into a maximal clique Q and an independent set I . Suppose, for the sake of contradiction, that Q contains a vertex u with two neighbors $a, b \in I$. As we took Q to be a maximal clique, a has a nonneighbor v and b has a nonneighbor w in Q . If a, w are not adjacent, then the vertices a, b, u, w induce a claw in G ; if b, v are not adjacent, then the vertices a, b, u, v induce a claw in G ; otherwise the vertices a, b, u, v, w induce a 3-fan in G . In either case we get a contradiction.

(d) $\mathcal{K}_4 = \text{Free}(\mathcal{F})$ for $\mathcal{F} = \{2K_2, C_4, C_5, K_4^-\}$: Again, it is easy to check that

none of the forbidden graphs belong to \mathcal{K}_4 . Let $G \in \text{Free}(\mathcal{F})$. By Theorem 3.2, G is a split graph. Just like before, split G into a maximal clique Q and an independent set I . Suppose that some vertex u in I has two neighbors a, b in Q . By maximality of Q , u also has a nonneighbor c in Q . But then the vertices a, b, c, u induce a K_4^- in G , a contradiction.

(e) The class $\overline{\mathcal{K}_5}$ of the complements of the graphs in \mathcal{K}_5 is characterized as the class of all (bipartite) graphs whose vertex set can be split into independent sets I_1, I_2 so that each vertex in I_2 has at most one nonneighbor in I_1 . We show that $\overline{\mathcal{K}_5} = \text{Free}(\mathcal{F})$ for $\mathcal{F} = \{K_3, C_5, P_4 + K_1, 2K_2 + K_1, C_4 + K_2, C_4 + 2K_1, H_6\}$. The reader will kindly check that indeed no graph in \mathcal{F} belongs to $\overline{\mathcal{K}_5}$.

Consider some $G \in \text{Free}(\mathcal{F})$; we will show that $G \in \overline{\mathcal{K}_5}$. Observe that \mathcal{F} prevents G from having an odd cycle, thus G is bipartite. We distinguish three cases depending on the structure of the connected components of G .

First, suppose that G has at least two nontrivial connected components (i.e., connected components that are not just isolated vertices). Because G is $(2K_2 + K_1)$ -free, it only has two connected components in all. Being C_4 - and P_4 -free, each component is necessarily a star. Observe that any graph consisting of one or two stars belongs to $\overline{\mathcal{K}_5}$.

Next assume that G has only one nontrivial connected component and some isolated vertices. The nontrivial component is bipartite and P_4 -free, so it is a biclique. If this biclique contains C_4 , then G only contains one other isolated vertex; any graph consisting of a biclique and one isolated vertex is in $\overline{\mathcal{K}_5}$. Otherwise the biclique is a star; any graph consisting of a star and one or more isolated vertices belongs to $\overline{\mathcal{K}_5}$.

Finally, consider G that is connected. We will show that for any two vertices of G in different parts, one of them must have at most one nonneighbor in the opposite part. Suppose this is not true and there are $x, y \in V(G)$ in different parts such that both x and y have at least two nonneighbors in the opposite part. Assume first that x and y are adjacent. Let a and b be two nonneighbors of x , and let c and d be two nonneighbors of y . Then the graph induced by a, b, c , and d cannot be a C_4 , P_4 , $P_3 + K_1$, $2K_2$, or $K_2 + 2K_1$, because G is $(P_4 + K_1, 2K_2 + K_1, C_4 + K_2)$ -free. Hence a, b, c , and d must induce a $4K_1$. As G is connected, a must have a neighbor, say w . However, the vertices x, y, a, c , and w induce a $P_4 + K_1$ if y and w are adjacent and they induce a $2K_2 + K_1$ if y and w are not adjacent. Therefore, x and y must be nonneighbors.

By assumption, x has another nonneighbor $a \neq y$ in the opposite part, and y has another nonneighbor $b \neq x$ in the opposite part. As G is connected, x must have a neighbor, say u . If a and b are adjacent, then x, y, u, a , and b induce a $2K_2 + K_1$ if u is not adjacent to b , and they induce a $P_4 + K_1$ if u is adjacent to b . Both cases lead to a contradiction as G is $(P_4 + K_1, 2K_2 + K_1)$ -free, hence a and b cannot be adjacent. Now, as G is connected, y must also have a neighbor, say v . If u is not adjacent to b , then x, y, u, v , and b induce either a $2K_2 + K_1$ or a $P_4 + K_1$, hence u and b must be adjacent. By a symmetric argument, v is adjacent to a . Now u and v must be nonadjacent, otherwise x, y, u, v, a , and b induce an H_6 .

This argument shows that any neighbor of x must also be a neighbor of b , any neighbor of y must also be a neighbor of a , and that any neighbor of x cannot be adjacent to any neighbor of y . This means that the shortest induced path between x and y must contain a P_6 , which is a contradiction as G is $(P_4 + K_1)$ -free. Therefore, either x or y must have at most one nonneighbor. This implies that G can be split into two independent sets I_1, I_2 such that every vertex in I_2 has at most one nonneighbor in I_1 , so G belongs to $\overline{\mathcal{K}_5}$.

(f) Chain graphs are characterised by finitely many forbidden induced subgraphs by a result of Yannakakis [18]; namely, $\mathcal{K}_6 = \text{Free}(2K_2, K_3, C_5)$.

(g) Threshold graphs are characterized by finitely many forbidden induced subgraphs by a result of Chvátal and Hammer [8]; namely, $\mathcal{K}_7 = \text{Free}(2K_2, P_4, C_4)$. \square

3.2. Finite distinguishing number. In this section we provide a characterization of the minimal classes for the case of finite distinguishing number $k_{\mathcal{X}}$. It turns out that these minimal classes consist of $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graphs, that is, the vertex set of each graph is partitioned into at most $\ell_{\mathcal{X}}$ bags and dense pairs are defined by a density graph H (see Lemma 2.11). The condition of Theorem 2.12 is enforced by long paths (indeed, an infinite path in the infinite universal graph). Thus actually $d_{\mathcal{X}} \leq 2$ for the minimal classes \mathcal{X} .

Let A be a finite alphabet. A *word* is a mapping $w : S \rightarrow A$, where $S = \{1, 2, \dots, n\}$ for some $n \in \mathbb{N}$ or $S = \mathbb{N}$; $|S|$ is the *length* of w , denoted by $|w|$. We write w_i for $w(i)$, and we often use the notation $w = w_1w_2w_3 \dots w_n$ or $w = w_1w_2w_3 \dots$. For $n \leq m$ and $w = w_1w_2 \dots w_n$, $w' = w'_1w'_2 \dots w'_m$ (or $w' = w'_1w'_2 \dots$), we say that w is a *factor* of w' if there exists a nonnegative integer s such that $w_i = w'_{i+s}$ for $1 \leq i \leq n$; w is an *initial segment* of w' if we can take $s = 0$.

Let H be an undirected graph with loops allowed and with vertex set $V(H) = A$, and let w be a (finite or infinite) word over the alphabet A . For any increasing sequence $u_1 < u_2 < \dots < u_m$ of positive integers such that $u_m \leq |w|$, define $G_{w,H}(u_1, u_2, \dots, u_m)$ to be the graph with vertex set $\{u_1, u_2, \dots, u_m\}$ and an edge between u_i and u_j if and only if

- either $|u_i - u_j| = 1$ and $w_{u_i}w_{u_j} \notin E(H)$
- or $|u_i - u_j| > 1$ and $w_{u_i}w_{u_j} \in E(H)$.

Let $G = G_{w,H}(u_1, u_2, \dots, u_m)$ and define $V_a = \{u_i \in V(G) : w_{u_i} = a\}$ for any $a \in A$. Then $\pi = \pi_w(G) = \{V_a : a \in A\}$ is an $(|A|, 2)$ -partition, and so G is an $(|A|, 2)$ -graph. Moreover, $\psi(G, \pi, H)$ is a linear forest whose paths are formed by the segments of consecutive integers within the set $\{u_1, u_2, \dots, u_m\}$. This partition $\pi_w(G)$ is called the *letter partition* of G .

DEFINITION 3.4. Let H be an undirected graph with loops allowed and with vertex set $V(H) = A$, and let w be an infinite word over the alphabet A . Define $\mathcal{P}(w, H)$ to be the hereditary class consisting of the graphs $G_{w,H}(u_1, u_2, \dots, u_m)$ for all finite increasing sequences $u_1 < u_2 < \dots < u_m$ of positive integers.

As we shall see later, all classes $\mathcal{P}(w, H)$ are above the Bell number. More importantly, all minimal classes above the Bell number have the form $\mathcal{P}(w, H)$ for some w and H . Our goal here is first to describe sufficient conditions on the word w under which $\mathcal{P}(w, H)$ is a minimal class above the Bell number; moreover, we aim to prove that any hereditary class above the Bell number with finite distinguishing number contains the class $\mathcal{P}(w, H)$ for some word w and graph H . We start by showing that these classes indeed have finite distinguishing number.

LEMMA 3.5. For any word w and graph H with loops allowed, the class $\mathcal{X} = \mathcal{P}(w, H)$ has finite distinguishing number.

Proof. Put $\ell = |H|$ and let G be a graph in \mathcal{X} . Consider the letter partition $\pi = \pi_w(G) = \{V_a : a \in V(H)\}$ of G , which is an $(\ell, 2)$ -partition. Choose an arbitrary set of vertices $X \subseteq V(G)$ and let $\{U_1, U_2, \dots, U_k\}$ be the sets distinguished by X . If there are subsets U_i, U_j , and V_a such that $|V_a \cap U_i| \geq 3$ and $|V_a \cap U_j| \geq 3$, then some vertex of X has at least three neighbors and at least three nonneighbors in V_a , which contradicts the fact that π is an $(\ell, 2)$ -partition. Therefore, in the partition

$\{V_a \cap U_i : a \in V(H), 1 \leq i \leq k\}$ we have at most ℓ sets of size at least 3. Note that every set U_i of size at least $2\ell + 1$ must contain at least one such set. Hence the family $\{U_1, U_2, \dots, U_k\}$ contains at most ℓ sets of size at least $2\ell + 1$. Since the set X was chosen arbitrarily, we conclude that $k_{\mathcal{X}} \leq 2\ell + 1$, as required. \square

The graphs $G_{w,H}(u_1, u_2, \dots, u_n)$ defined on a sequence of consecutive integers will play a special role in our considerations.

DEFINITION 3.6. *If u_1, u_2, \dots, u_m is a sequence of consecutive integers (i.e., $u_{k+1} = u_k + 1$ for each k), we call the graph $G_{w,H}(u_1, u_2, \dots, u_m)$ an $|H|$ -factor. Notice that each $|H|$ -factor is an $(|H|, 2)$ -graph; if its letter partition is a strong $(|H|, 2)$ -partition, we call it a strong $|H|$ -factor.*

Note that if $G = G_{w,H}(u_1, u_2, \dots, u_m)$ is a strong $|H|$ -factor, then its sparsification $\phi(G) = \psi(G, \pi_w(G), H)$ is an induced path with m vertices.

PROPOSITION 3.7. *If w is an infinite word over a finite alphabet A and H is a graph on A , with loops allowed, then the class $\mathcal{P}(w, H)$ is above the Bell number.*

Proof. We may assume that every letter of A appears in w infinitely many times; otherwise, we can remove a sufficiently long starting segment of w to obtain a word w' satisfying this condition, replace H with its induced subgraph H' on the alphabet A' of w' , and obtain a subclass $\mathcal{P}(w', H')$ of $\mathcal{P}(w, H)$ with that property. For sufficiently large k , the $|A|$ -factor $G_k = G_{w,H}(1, \dots, k)$ is a strong $|A|$ -factor; thus $\phi(G_k)$ is an induced path of length $k - 1$. Having a finite distinguishing number by Lemma 3.5, the class $\mathcal{P}(w, H)$ is above the Bell number by Theorem 2.12. \square

DEFINITION 3.8. *An infinite word w is called almost periodic if for any factor f of w there is a constant k_f such that any factor of w of length at least k_f contains f as a factor.*

The notion of an almost periodic word plays a crucial role in our characterization of minimal classes above the Bell number. First, let us show that if w is almost periodic, then $\mathcal{P}(w, H)$ is a minimal property above the Bell number. To prove this, we need an auxiliary lemma.

LEMMA 3.9. *Consider $G = G_{w,H}(u_1, \dots, u_n)$. If G is a strong (ℓ, d) -graph and $\phi(G)$ contains a connected component C such that $|C| \geq (2d^2\ell^2|H|^2 + 1)(m - 1) + 1$, where $d' = \max\{d, 2\}$, then $V(C)$ contains a sequence of m consecutive integers.*

Proof. Let $\pi = \{U_1, U_2, \dots, U_{\ell'}\}$ be a strong (ℓ, d) -partition of G , so that $\ell' \leq \ell$ and $\phi(G) = \phi(G, \pi)$; let $\pi' = \{V_a : a \in V(H)\}$ be the letter partition of G , given by $V_a = \{u_j \in V(G) : w_{u_j} = a\}$. Put $k = |H|$. Note that π' is a $(k, 2)$ -partition, hence also a (k, d') -partition.

Let $E = E(\phi(G)) \setminus E(\psi(G, \pi', H))$ be the set of all the edges of $\phi(G)$ that are not edges of $\psi(G, \pi', H)$, that is, that do not join two consecutive integers. We will now upper bound the number of such edges. Observe that E consists of (a) the edges between $U_i \cap V_a$ and $U_j \cap V_b$, where (U_i, U_j) is d' -sparse and (V_a, V_b) is d' -dense, and (b) the nonedges between $U_i \cap V_a$ and $U_j \cap V_b$, where (U_i, U_j) is d' -dense and (V_a, V_b) is d' -sparse.

Consider the partition $\rho = \{U_i \cap V_a : 1 \leq i \leq \ell', a \in V(H)\}$ of G , which is an $(\ell'k, d')$ -partition. Let $(U_i \cap V_a, U_j \cap V_b)$ be a pair of nonempty sets such that (U_i, U_j) is d' -sparse but (V_a, V_b) is d' -dense. Each such pair is both d' -sparse and d' -dense, and consequently we have $|U_i \cap V_a| \leq 2d'$ and $|U_j \cap V_b| \leq 2d'$. Moreover, there are at most $2d'^2$ edges between $U_i \cap V_a$ and $U_j \cap V_b$. Similarly, for any pair $(U_i \cap V_a, U_j \cap V_b)$,

where (U_i, U_j) is d' -dense but (V_a, V_b) is d' -sparse, we can show that there are at most $2d'^2$ nonedges between $U_i \cap V_a$ and $U_j \cap V_b$. We conclude that $|E| \leq 2d'^2(\ell'k)^2$.

Any edge of $\phi(G)$ that is not in E joins two consecutive integers. Hence any connected component C of $\phi(G)$ consists of at most $|E| + 1$ segments of consecutive integers connected by edges from E . If C does not contain a sequence of m consecutive integers, it consists of at most $|E| + 1 \leq 2d'^2(\ell'k)^2 + 1$ segments of consecutive integers, each of length at most $m - 1$; it can therefore contain at most $(2d'^2(\ell'k)^2 + 1)(m - 1) \leq (2d'^2\ell^2|H|^2 + 1)(m - 1)$ vertices. \square

THEOREM 3.10. *If w is an almost periodic infinite word and H is a finite graph with loops allowed, then $\mathcal{P}(w, H)$ is a minimal hereditary property above the Bell number.*

Proof. The class $\mathcal{P} = \mathcal{P}(w, H)$ is above the Bell number by Proposition 3.7. Thus we only need to show that any proper hereditary subclass \mathcal{X} of \mathcal{P} is below the Bell number. Suppose $\mathcal{X} \subset \mathcal{P}$ and let $F \in \mathcal{P} \setminus \mathcal{X}$. By the definition of $\mathcal{P}(w, H)$, the graph F is of the form $G_{w,H}(u_1, \dots, u_n)$ for some positive integers $u_1 < \dots < u_n$. Let w' be the finite word $w' = w_{u_1}w_{u_1+1}w_{u_1+2} \dots w_{u_n-1}w_{u_n}$. As w is almost periodic, there is an integer m such that any factor of w of length m contains w' as a factor. Assume, for the sake of contradiction, that \mathcal{X} is hereditary and above the Bell number. By Lemma 3.5, the distinguishing number of \mathcal{P} , and hence of \mathcal{X} , is finite, and therefore, by Lemma 2.11 and Theorem 2.12, there exists a strong $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph $G = G_{w,H}(u'_1, u'_2, \dots, u'_{n'}) \in \mathcal{X}$ such that $\phi(G)$ has a connected component C of order at least $(2d^2\ell^2|H|^2 + 1)(m - 1) + 1$, where $\ell = \ell_{\mathcal{X}}$ and $d = \max\{d_{\mathcal{X}}, 2\}$. By Lemma 3.9, the vertices of C contain a sequence of m consecutive integers, i.e., $V(G) \supseteq V(C) \supseteq \{u', u' + 1, \dots, u' + m - 1\}$. However, the word $w_{u'}w_{u'+1} \dots w_{u'+m-1}$ contains w' ; therefore G contains F , a contradiction. \square

The existence of minimal classes does not necessarily imply that every class above the Bell number contains a minimal one. However, in our case this turns out to be true, as we proceed to show next. Moreover, this will also imply that the minimal classes described in Theorem 3.10 are the *only* minimal classes above the Bell number with $k_{\mathcal{X}}$ finite. To prove this, we first show in the next two lemmas that any class \mathcal{X} above the Bell number with $k_{\mathcal{X}}$ finite contains arbitrarily large strong $\ell_{\mathcal{X}}$ -factors.

LEMMA 3.11. *Let \mathcal{X} be a hereditary class with speed above the Bell number and with finite distinguishing number $k_{\mathcal{X}}$. Then for each m , the class \mathcal{X} contains an $\ell_{\mathcal{X}}$ -factor of order m .*

Proof. From Theorem 2.12 it follows that for each m there is a graph $G_m \in \mathcal{X}$ which admits a strong $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -partition $\{V_1, V_2, \dots, V_{\ell_m}\}$ with $\ell_m \leq \ell_{\mathcal{X}}$ such that the sparsification $\phi(G_m)$ has a connected component C_m of order at least $(\ell_{\mathcal{X}}d_{\mathcal{X}})^m$. Fix an arbitrary vertex v of C_m . As C_m is an induced subgraph of $\phi(G_m)$, the maximum degree in C_m is bounded by $d = \ell_{\mathcal{X}}d_{\mathcal{X}}$. Hence for any $k > 0$, in C_m there are at most $d(d - 1)^{k-1}$ vertices at distance k from v ; so there are at most $1 + \sum_{k=1}^{m-2} d(d - 1)^{k-1} < d^m$ vertices at distance at most $m - 2$ from v . As C_m has order at least d^m , there exists a vertex v' of distance $m - 1$ from v . Therefore C_m contains an induced path $v = v_1, v_2, \dots, v_m = v'$ of length $m - 1$.

Let $A = \{1, 2, \dots, \ell_m\}$ and let H be the graph with vertex set A and edge between i and j if and only if V_i is $d_{\mathcal{X}}$ -dense with respect to V_j . Let $w_i \in A$ be such that $v_i \in V_{w_i}$ and define the word $w = w_1w_2 \dots w_m$. The induced subgraph $G_m[v_1, v_2, \dots, v_m] \cong G_{w,H}(1, 2, \dots, m)$ is an $\ell_{\mathcal{X}}$ -factor of order m contained in \mathcal{X} . \square

LEMMA 3.12. *Let ℓ and B be positive integers such that $B \geq 5 \times 2^{\ell+1}$. Then we have that any ℓ -factor $G_{w,H}(1, 2, \dots, |w|)$ of order at least B^ℓ contains a strong ℓ -factor $G_{w',H}(1, 2, \dots, |w'|)$ of order at least B such that w' is a factor of w .*

Proof. To prove this we will show by induction on $r \in \{1, 2, \dots, \ell\}$ that any ℓ -factor $G_{w,H}(1, 2, \dots, B^r)$ on B^r vertices with at most r bags in the letter partition contains a strong ℓ -factor on at least B vertices. For $r = 1$ the statement holds because any ℓ -factor with one bag in the letter partition of order $B \geq 5 \times 2^{\ell+1}$ is a strong ℓ -factor. Suppose $1 < r \leq \ell$. Then either each letter of $w = w_1 w_2 \dots w_{B^r}$ appears at least B times, in which case we are done, or there is a letter $a = w_i$ which appears less than B times in w . Consider the maximal factors of w that do not contain the letter a . Because the number of occurrences of the letter a in w is less than B , there are at most B such factors of w and the sum of their orders is at least $B^r - B + 1$. By the pigeonhole principle, one of these factors has order at least B^{r-1} ; call this factor w'' . Now w'' contains at most $r - 1$ different letters; thus $G'' = G_{w'',H}(1, 2, \dots, |w''|)$ is an ℓ -factor of order at least B^{r-1} for which the letter partition has at most $(r - 1)$ bags. By induction, G'' contains a strong ℓ -factor $G_{w',H}(1, 2, \dots, |w'|)$ of order at least B such that w' is a factor of w'' which is a factor of w . Hence w' is a factor of w and we are done. \square

THEOREM 3.13. *Suppose \mathcal{X} is a hereditary class above the Bell number with $k_{\mathcal{X}}$ finite. Then $\mathcal{X} \supseteq \mathcal{P}(w, H)$ for an infinite almost periodic word w and a graph H of order at most $\ell_{\mathcal{X}}$ with loops allowed.*

Proof. From Lemmas 3.11 and 3.12 it follows that each class \mathcal{X} with speed above the Bell number with finite distinguishing number $k_{\mathcal{X}}$ contains an infinite set \mathcal{S} of strong $\ell_{\mathcal{X}}$ -factors of increasing order. For each H on $\{1, 2, \dots, \ell\}$ with $1 \leq \ell \leq \ell_{\mathcal{X}}$, let $\mathcal{S}_H = \{G_{w,H}(1, \dots, |w|) \in \mathcal{S}\}$ be the set of all $\ell_{\mathcal{X}}$ -factors in \mathcal{S} whose adjacencies are defined using the density graph H . Then for some (at least one) fixed graph H_0 the set \mathcal{S}_{H_0} is infinite. Hence also $L = \{w : G_{w,H_0}(1, \dots, |w|) \in \mathcal{X}\}$ is an infinite language. As \mathcal{X} is a hereditary class, the language L is closed under taking word factors (it is a *factorial language*).

It is not hard to see that any infinite factorial language contains an inclusion-minimal infinite factorial language. So let $L' \subseteq L$ be a minimal infinite factorial language. It follows from the minimality that L' is well quasi-ordered by the factor relation, because otherwise removing one word from any infinite antichain and taking all factors of the remaining words would generate an infinite factorial language strictly contained in L' . Thus there exists an infinite chain $w^{(1)}, w^{(2)}, \dots$ of words in L' such that for any $i < j$, the word $w^{(i)}$ is a factor of $w^{(j)}$. More precisely, for each i there is a nonnegative integer s_i such that $w_k^{(i)} = w_{k+s_i}^{(i+1)}$. Let $g(i, k) = k + \sum_{j=1}^{i-1} s_j$. Now we can define an infinite word w by putting $w_k = w_{g(i,k)}^{(i)}$ for the least value of i for which the right-hand side is defined. (Without loss of generality we get that w is indeed an infinite word; otherwise we would need to take the reversals of all the words $w^{(i)}$.)

Observe that any factor of w is a factor of some $w^{(i)}$ and hence in the language L' . If w is not almost periodic, then there exists a factor f of w such that there are arbitrarily long factors f' of w not containing f . These factors f' generate an infinite factorial language $L'' \subset L'$ which does not contain $f \in L'$. This contradicts the minimality of L' and proves that w is almost periodic.

Because any factor of w is in L , any $G_{w,H_0}(u_1, \dots, u_m)$ is an induced subgraph of some $\ell_{\mathcal{X}}$ -factor in \mathcal{X} . Therefore $\mathcal{P}(w, H_0) \subseteq \mathcal{X}$. \square

Combining Theorems 3.10 and 3.13 we derive the main result of this section.

COROLLARY 3.14. *Let \mathcal{X} be a class of graphs with $k_{\mathcal{X}} < \infty$. Then \mathcal{X} is a minimal hereditary class above the Bell number if and only if there exists a finite graph H with loops allowed and an infinite almost periodic word w over $V(H)$ such that $\mathcal{X} = \mathcal{P}(w, H)$.*

Last, note that—similarly to the case of an infinite distinguishing number—each of the minimal classes $\mathcal{P}(w, H)$ has an infinite universal graph $G_{w,H}(1, 2, 3, \dots)$.

4. Decidability of the Bell number. As we mentioned in the introduction, every class below the Bell number can be characterized by a finite set of forbidden induced subgraphs. Therefore all classes for which the set of minimal forbidden induced subgraphs is infinite have speed above the Bell number. For classes defined by finitely many forbidden induced subgraphs, the problem of deciding whether their speed is above the Bell number is more complicated and decidability of this problem has been an open question. In this section, we employ our characterization of minimal classes above the Bell number to answer this question positively.

Our main goal is to provide an algorithm that decides, for an input consisting of a finite number of graphs F_1, \dots, F_n , whether the speed of $\mathcal{X} = \text{Free}(F_1, \dots, F_n)$ is above the Bell number. That is, we are interested in the following problem.

PROBLEM 4.1.

INPUT: *A finite set of graphs $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$.*

OUTPUT: *Yes, if the speed of $\mathcal{X} = \text{Free}(\mathcal{F})$ is above the Bell number; no otherwise.*

Our algorithm, following the characterization of minimal classes above the Bell number, distinguishes two cases depending on whether the distinguishing number $k_{\mathcal{X}}$ is finite or infinite. First we show how to discriminate between these two cases.

PROBLEM 4.2.

INPUT: *A finite set of graphs $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$.*

OUTPUT: *Yes, if $k_{\mathcal{X}} = \infty$ for $\mathcal{X} = \text{Free}(\mathcal{F})$; no otherwise.*

THEOREM 4.3. *There is a polynomial-time algorithm that solves Problem 4.2.*

Proof. By Theorem 3.1, $k_{\mathcal{X}} = \infty$ if and only if \mathcal{X} contains one of the thirteen minimal classes listed there. By Theorem 3.3, each of the minimal classes is defined by finitely many forbidden induced subgraphs; thus membership can be tested in polynomial time. Then the answer to Problem 4.2 is no if and only if each of the minimal classes given by Theorem 3.1 contains at least one of the graphs in \mathcal{F} , which can also be tested in polynomial time. \square

By Corollary 3.14, the minimal hereditary classes with finite distinguishing number with speed above the Bell number can be described as $\mathcal{P}(w, H)$ with an almost periodic infinite word w . That characterization applies both to classes defined by finitely many forbidden subgraphs and to classes defined by infinitely many forbidden subgraphs. In the case of finitely many forbidden subgraphs, however, a stronger characterization is possible, as we show next.

DEFINITION 4.4. *Let $w = w_1w_2\dots$ be an infinite word over a finite alphabet A . If there exists some p such that $w_i = w_{i+p}$ for all $i \in \mathbb{N}$, we call the word w periodic and the number p its period. If, moreover, for some period p the letters w_1, w_2, \dots, w_p are all distinct, we call the word w cyclic.*

If w is a finite word, then w^∞ is the periodic word obtained by concatenating infinitely many copies of the word w ; thus $(w^\infty)_i = w_k$ for $k = i \bmod |w|$.

A class \mathcal{X} of graphs is called a periodic class (cyclic class, respectively) if there exists a graph H with loops allowed and a periodic (cyclic, respectively) word w such that $\mathcal{X} = \mathcal{P}(w, H)$.

DEFINITION 4.5. Let $A = \{1, 2, \dots, \ell\}$ be a finite alphabet, H a graph on A with loops allowed, and m a positive integer. Define a graph $S_{H,m}$ with vertex set $V(S_{H,m}) = A \times \{1, 2, \dots, m\}$ and an edge between (a, j) and (b, k) if and only if $(a, j) \neq (b, k)$ and one of the following holds:

- $ab \in E(H)$ and either $|a - b| \neq 1$ or $j \neq k$ (or both);
- $ab \notin E(H)$ and $|a - b| = 1$ and $j = k$.

The graph $S_{H,m}$ is called an (ℓ, m) -strip.

Notice that a strip can be viewed as the graph obtained from the union of m disjoint paths $(1, j) - (2, j) - \dots - (\ell, j)$ for $j \in \{1, 2, \dots, m\}$ by swapping edges with nonedges between vertices (a, j) and (b, k) if $ab \in E(H)$.

THEOREM 4.6. Let $\mathcal{X} = \text{Free}(F_1, F_2, \dots, F_n)$ with $k_{\mathcal{X}}$ finite. Then the following conditions are equivalent:

- (a) The speed of \mathcal{X} is above the Bell number.
- (b) \mathcal{X} contains a periodic class.
- (c) For every $p \in \mathbb{N}$, \mathcal{X} contains a cyclic class with period at least p .
- (d) There exists a cyclic word w and a graph H on the alphabet of w such that \mathcal{X} contains the ℓ -factor $G_{w,H}(1, 2, \dots, 2\ell m)$ with $\ell = |V(H)|$ and also with $m = \max\{|F_i| : i \in \{1, 2, \dots, n\}\}$.
- (e) For any positive integers ℓ, m , the class \mathcal{X} contains an (ℓ, m) -strip.

Proof. (a) \Rightarrow (b): From Theorem 3.13 we know that \mathcal{X} contains a class $\mathcal{P}(w, H)$ with some almost periodic word $w = w_1 w_2 \dots$ and a finite graph H with loops allowed. Let $m = \max\{|F_1|, |F_2|, \dots, |F_n|\}$ and let $a = w_1 w_2 \dots w_m$ be the word consisting of the first m letters of the infinite word w . Since w is almost periodic, the factor a appears in w infinitely often. In particular, there is $m' > m$ such that $w_{m'+1} w_{m'+2} \dots w_{m'+m} = a$. Define b to be the word between the two a 's in w , i.e., let $b = w_{m+1} w_{m+2} \dots w_{m'}$. In this way, w starts with the initial segment aba .

We claim that \mathcal{X} contains the periodic class $\mathcal{P}(w', H)$ with $w' = (ab)^\infty$. For the sake of contradiction, suppose that \mathcal{X} does not contain $\mathcal{P}(w', H)$. Then for some $i \in \{1, 2, \dots, n\}$, we have $F_i \in \mathcal{P}(w', H)$. So $F_i \cong G_{w',H}(u_1, u_2, \dots, u_k)$ for some $u_1 < u_2 < \dots < u_k$. Let $U = \{u_1, u_2, \dots, u_k\}$. We are now looking for a monotonically increasing function $f : U \rightarrow \mathbb{N}$ with these two properties: first, $w_{f(u)} = w'_u$ for any $u \in U$; second, $f(u) - f(u') = 1$ if and only if $u - u' = 1$. If we can establish the existence of such a function, we will then have $F_i \cong G_{w',H}(u_1, u_2, \dots, u_k) \cong G_{w,H}(f(u_1), f(u_2), \dots, f(u_k)) \in \mathcal{P}(w, H) \subseteq \mathcal{X}$, a contradiction.

To construct such a function f , consider a maximal block $\{u_j, u_{j+1}, \dots, u_{j+p}\}$ of consecutive integers in U (that is, $u_{j-1} < u_j - 1$; $u_j = u_{j+1} - 1$; \dots ; $u_{j+p-1} = u_{j+p} - 1$; $u_{j+p} < u_{j+p+1} - 1$). Furthermore, consider the word $w'_{u_j} w'_{u_{j+1}} \dots w'_{u_{j+p}}$ of length at most m , which is a factor of $w' = (ab)^\infty$ and thus also a factor of aba because $|aba| > 2m$. The word aba , being a factor of w , appears infinitely often in w because w is almost periodic. Hence not only we can define $f : U \rightarrow \mathbb{N}$ in such a way that $w_{f(u)} = w'_u$ for any $u \in U$ and that blocks of consecutive integers in U are mapped to blocks of consecutive integers, but we can also do it monotonically and so that $f(u) > f(u') + 1$ whenever $u > u' + 1$. This finishes the proof of the first implication.

(b) \Rightarrow (c): Let \mathcal{X} contain a class $\mathcal{P}(w, H)$, where w is a periodic word and H is a graph with loops allowed on the alphabet of w . If w is not cyclic (i.e., if some letters appear more than once within the period), the class $\mathcal{P}(w, H)$ can be transformed into a cyclic one by extending the alphabet, renaming multiple appearances of the same letter within the period to different letters of the extended alphabet, and modifying the graph H accordingly. More formally, let p be the period of w . Define a new word $w' = (123 \dots p)^\infty$ and a graph H' with vertex set $\{1, 2, \dots, p\}$ and an edge $ij \in E(H')$ if and only if $w_i w_j \in E(H)$. Then any graph $G_{w,H}(u_1, u_2, \dots, u_m) \in \mathcal{P}(w, H)$ is isomorphic to the graph $G_{w',H'}(u_1, u_2, \dots, u_m) \in \mathcal{P}(w', H')$. Hence \mathcal{X} also contains $\mathcal{P}(w', H')$, where the word w' is cyclic.

Since any periodic word of period p is also a periodic word of period kp for any $k \in \mathbb{N}$, by the same transformation a periodic class of period p can be transformed into a cyclic class of period kp .

(c) \Rightarrow (d): Follows directly from the definition of a cyclic class.

(d) \Rightarrow (a): Let \mathcal{X} contain the ℓ -factor $G_{w,H}(1, 2, \dots, 2\ell m)$. We prove that \mathcal{X} then contains the whole class $\mathcal{P}(w, H)$; its speed will then be above the Bell number by Proposition 3.7. For the sake of contradiction suppose that for some i , the graph F_i belongs to $\mathcal{P}(w, H)$. Then $F_i \cong G_{w,H}(u_1, u_2, \dots, u_k)$ for $k = |F_i| \leq m$ and $u_1 < u_2 < \dots < u_k$. Note that the period of w is ℓ (by the definition of a cyclic class). Put $u'_1 = u_1 \bmod \ell$, or $u'_1 = \ell$ if $u_1 \equiv 0 \pmod{\ell}$. Furthermore, for each $i \geq 2$ let $u'_i = (u_i \bmod \ell) + c_i \ell$, where each c_i is chosen in such a way that $0 < u'_i - u'_{i-1} \leq \ell + 1$ for all i and $u'_i - u'_{i-1} = 1$ if and only if $u_i - u_{i-1} = 1$. By construction, $F_i \cong G_{w,H}(u_1, u_2, \dots, u_k) \cong G_{w,H}(u'_1, u'_2, \dots, u'_k)$ and $u'_k < k(\ell + 1) \leq m(\ell + 1) \leq 2\ell m$. Hence F_i is isomorphic to an induced subgraph of $G_{w,H}(1, 2, \dots, 2\ell m) \in \mathcal{X} = \text{Free}(F_1, F_2, \dots, F_n)$, a contradiction.

(c) \Rightarrow (e): Let \mathcal{X} contain the cyclic class $\mathcal{P}(w, H)$, where w is a cyclic word with period $p > \ell$. Since $p > \ell$, the subgraph of $G_{w,H}(1, 2, \dots, pm)$ induced by the bags corresponding to the first ℓ letters of w is an (ℓ, m) -strip.

(e) \Rightarrow (a): Let $\ell_{\mathcal{X}}$, $d_{\mathcal{X}}$, and $c_{\mathcal{X}}$ be the constants given by Lemma 2.11 and put $m = 2d_{\mathcal{X}} + c_{\mathcal{X}} + 2$. We show that for any fixed positive integer ℓ , the class \mathcal{X} contains a strong $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph G such that its sparsification $\phi(G)$ has a connected component of order at least ℓ . Then we can apply Theorem 2.12.

So let ℓ be a positive integer. By assumption, \mathcal{X} contains an (ℓ, m) -strip $S_{H,m}$. By Lemma 2.11, after removing no more than $c_{\mathcal{X}}$ vertices we are left with a strong $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph S' with a strong $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -partition π . Then, we will let $V_a = \{(a, j) \in V(S') : 1 \leq j \leq m\}$ be the letter bags of S' , $1 \leq a \leq \ell$, and consider the prime partition $p(\pi) = \{W_1, W_2, \dots, W_\ell\}$. If two vertices x, y belong to different bags of $p(\pi)$, then according to Lemma 2.5 we have $|N(x) \ominus N(y)| \geq 5 \times 2^{\ell_{\mathcal{X}}} d_{\mathcal{X}} - 2d_{\mathcal{X}} \geq 8$. However, if we have two vertices $(a, j), (a, j')$ of S' in the same letter bag V_a , then $N(a, j) \ominus N(a, j') \subseteq \{(a-1, j), (a-1, j'), (a, j), (a, j'), (a+1, j), (a+1, j')\}$, so its size is at most 6. Hence, we deduce that each $V_a \subseteq W_{f(a)}$ for some function f .

Now notice that (V_a, V_b) is $d_{\mathcal{X}}$ -dense, that is, ab is an edge of H , if and only if $(W_{f(a)}, W_{f(b)})$ is $d_{\mathcal{X}}$ -dense. Indeed, if one of them was $d_{\mathcal{X}}$ -dense and the other $d_{\mathcal{X}}$ -sparse, then (V_a, V_b) would be both $d_{\mathcal{X}}$ -dense and $d_{\mathcal{X}}$ -sparse, in which case $|V_a| \leq 2d_{\mathcal{X}} + 1$. But this is not true, as V_a is obtained from a set of size $m = 2d_{\mathcal{X}} + 2 + c_{\mathcal{X}}$ by removing at most $c_{\mathcal{X}}$ vertices.

It follows that $\phi(S')$ is constructed by swapping edges with nonedges between V_a and V_b such that $ab \in E(H)$. Hence $\phi(S')$ is a linear forest obtained from the paths $(1, j) - (2, j) - \dots - (\ell, j)$ for $j \in \{1, 2, \dots, m\}$ by removing at most $c_{\mathcal{X}}$ vertices.

As $m > c_{\mathcal{X}}$, at least one of the paths is left untouched. Therefore, $\phi(S')$ contains a connected component of size at least ℓ . \square

Finally, we are ready to tackle the decidability of Problem 4.1.

ALGORITHM 4.7.

INPUT: A finite set of graphs $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$.

OUTPUT: Yes, if the speed of $\mathcal{X} = \text{Free}(\mathcal{F})$ is above the Bell number; no otherwise.

(1) Using Theorem 4.3, decide whether $k_{\mathcal{X}} = \infty$. If it is, output yes and stop.

(2) Set $m := \max\{|F_1|, |F_2|, \dots, |F_n|\}$ and $\ell := 1$.

(3) Loop:

(3a) For each graph (with loops allowed) H on $\{1, 2, \dots, \ell\}$ construct the (ℓ, ℓ) -strip $S_{H, \ell}$. Check if some F_i is an induced subgraph of $S_{H, \ell}$. If for each H the strip $S_{H, \ell}$ contains some F_i , output no and stop.

(3b) For each graph (with loops allowed) H on $\{1, 2, \dots, \ell\}$ and for each word w consisting of ℓ distinct letters from $\{1, 2, \dots, \ell\}$ check to see if the ℓ -factor $G_{w^\infty, H}(1, 2, \dots, 2\ell m)$ contains some F_i as an induced subgraph. If one of these ℓ -factors contains no F_i , output yes and stop.

(3c) Set $\ell := \ell + 1$ and repeat.

It remains to prove the correctness of this algorithm.

THEOREM 4.8. Algorithm 4.7 correctly solves Problem 4.1.

Proof. We show that if the algorithm stops, it gives the correct answer, and furthermore that it will stop on any input without entering an infinite loop. First, if it stops in step (1), the answer is correct by [7], since any class with infinite distinguishing number has speed above the Bell number.

Assume that the algorithm stops in step (3a) and outputs *no*. This is because every (ℓ, ℓ) -strip contains some forbidden subgraph F_i , hence no (ℓ, ℓ) -strip belongs to \mathcal{X} . By Theorem 4.6(e), the speed of \mathcal{X} is below the Bell number.

Next suppose that the algorithm stops in step (3b) and answers *yes*. Then \mathcal{X} contains the ℓ -factor $G_{w^\infty, H}(1, 2, \dots, 2\ell m)$, where w^∞ is a cyclic word. Hence by Theorem 4.6(d) the speed of \mathcal{X} is above the Bell number.

Finally, if $k_{\mathcal{X}} = \infty$ the algorithm stops in step (1). If $k_{\mathcal{X}} < \infty$ and the speed of \mathcal{X} is above the Bell number, then by Theorem 4.6(d) the algorithm will stop in step (3b). If, on the other hand, the speed of \mathcal{X} is below the Bell number, then by Theorem 4.6(e) there exist positive integers ℓ, M such that \mathcal{X} contains no (ℓ, M) -strip. Let $N = \max\{\ell, M\}$. Obviously, \mathcal{X} contains no (N, N) -strip, because any (N, N) -strip contains some (many) (ℓ, M) -strips as induced subgraphs and \mathcal{X} is hereditary. Therefore the algorithm will stop in step (3a) after finitely many steps. \square

5. Concluding remarks. In this paper, we have characterized all minimal hereditary classes of graphs whose speed is at least the Bell number B_n . This characterization allowed us to show that the problem of determining if the speed of a hereditary class \mathcal{X} defined by finitely many forbidden induced subgraphs is above or below the Bell number is decidable, i.e., there is an algorithm that gives a solution to this problem in a finite number of steps. However, the complexity of this algorithm, in terms of the input forbidden graphs, remains an open question. In particular, it would be interesting to determine if there is a polynomial bound on the minimum ℓ such that the input class \mathcal{X} contains an ℓ -factor as in Theorem 4.6(d) if it is above the Bell number, and it fails to contain any (ℓ, ℓ) -strip as in Theorem 4.6(e) if it is below.

REFERENCES

- [1] V. E. ALEKSEEV, *Range of values of entropy of hereditary classes of graphs*, Diskret. Mat., 4 (1992), pp. 148–157 (in Russian); Discrete Math. Appl., 3 (1993), pp. 191–199 (in English).
- [2] V. E. ALEKSEEV, *On lower layers of a lattice of hereditary classes of graphs*, Diskretn. Anal. Issled. Oper. Ser. 1, 4 (1997), pp. 3–12 (in Russian).
- [3] P. ALLEN, V. LOZIN, AND M. RAO, *Clique-width and the speed of hereditary properties*, Electron. J. Combin., 16 (2009), R35.
- [4] N. ALON, J. BALOGH, B. BOLLOBÁS, AND R. MORRIS, *The structure of almost all graphs in a hereditary property*, J. Combin. Theory Ser. B, 101 (2011), pp. 85–110.
- [5] J. BALOGH, B. BOLLOBÁS, AND D. WEINREICH, *The speed of hereditary properties of graphs*, J. Combin. Theory Ser. B, 79 (2000), pp. 131–156.
- [6] J. BALOGH, B. BOLLOBÁS, AND D. WEINREICH, *The penultimate rate of growth for graph properties*, European J. Combin., 22 (2001), pp. 277–289.
- [7] J. BALOGH, B. BOLLOBÁS, AND D. WEINREICH, *A jump to the Bell number for hereditary graph properties*, J. Combin. Theory Ser. B, 95 (2005), pp. 29–48.
- [8] V. CHVÁTAL AND P. L. HAMMER, *Aggregation of inequalities in integer programming*, in Studies in Integer Programming Ann. Discrete Math. 1., North-Holland, Amsterdam, 1977, pp. 175–162.
- [9] P. ERDŐS, P. FRANKL, AND V. RÖDL, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin., 2 (1986), pp. 113–121.
- [10] P. ERDŐS, D. J. KLEITMAN, AND B. L. ROTHSCCHILD, *Asymptotic enumeration of K_n -free graphs*, in Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, Atti dei Convegni Lincei 17., Accademia Nazionale dei Lincei, Rome, 1976, pp. 19–27.
- [11] S. FÖLDES AND P. L. HAMMER, *Split graphs*, in Proceedings of the Eighth Southeastern Conference on Combinatorics, Graph Theory and Computing (Louisiana State University, Baton Rouge, LA, 1977), Congr. Numer. 19, Utilitas Mathematica, Winnipeg, Canada, 1977, pp. 311–315.
- [12] PH. G. KOLAITIS, H. J. PRÖMEL, AND B. L. ROTHSCCHILD, *K_{l+1} -free graphs: asymptotic structure and a 0–1 law*, Trans. Amer. Math. Soc., 303 (1987), pp. 637–671.
- [13] N. KORPELAINE AND V. LOZIN, *Two forbidden induced subgraphs and well-quasi-ordering*, Discrete Math., 311 (2011), pp. 1813–1822.
- [14] H. J. PRÖMEL AND A. STEGER, *Excluding induced subgraphs: Quadrilaterals*, Random Structures Algorithms, 2 (1991), pp. 55–71.
- [15] H. J. PRÖMEL AND A. STEGER, *Excluding induced subgraphs. III. A general asymptotic*, Random Structures Algorithms, 3 (1992), pp. 19–31.
- [16] H. J. PRÖMEL AND A. STEGER, *Excluding induced subgraphs. II. Extremal graphs*, Discrete Appl. Math., 44 (1993), pp. 283–294.
- [17] E. R. SCHEINERMAN AND J. S. ZITO, *On the size of hereditary classes of graphs*, J. Combin. Theory Ser. B, 61 (1994), pp. 16–39.
- [18] M. YANNAKAKIS, *The complexity of the partial order dimension problem*, SIAM J. Algebraic Discrete Methods, 3 (1982), pp. 351–358.