When Trees Grow Low: Shrubs and Fast MSO$_1$

Robert Ganian[1], Petr Hliněný[2], Jaroslav Nešetřil[3], Jan Obdržálek[2],
Patrice Ossona de Mendez[4], and Reshma Ramadurai[2]

[1] Institute for Informatics, Goethe University, Frankfurt, Germany
`ganian@fi.muni.cz`
[2] Faculty of Informatics, Masaryk University, Brno, Czech Republic[⋆]
`{hlineny,obdrzalek,ramadur}@fi.muni.cz`
[3] Computer Science Inst. of Charles University (IUUK), Praha, Czech Republic[⋆][⋆⋆]
`nesetril@kam.ms.mff.cuni.cz`
[4] CNRS UMR 8557, École des Hautes Études en Sciences Sociales, Paris, France[⋆⋆]
`pom@ehess.fr`

Abstract. Recent characterization [9] of those graphs for which coloured
MSO$_2$ model checking is fast raised the interest in the graph invariant
called *tree-depth*. Looking for a similar characterization for (coloured)
MSO$_1$, we introduce the notion of *shrub-depth* of a graph class. To prove
that MSO$_1$ model checking is fast for classes of bounded shrub-depth,
we show that shrub-depth exactly characterizes the graph classes having
interpretation in coloured trees of bounded height. We also introduce
a common extension of cographs and of graphs with bounded shrub-
depth — *m-partite cographs* (still of bounded clique-width), which are
well quasi-ordered by the relation "is an induced subgraph of" and there-
fore allow polynomial time testing of hereditary properties.

1 Introduction

In this paper, we are interested in graph parameters that are intermediate be-
tween clique-width and tree-depth, sharing the nice properties of both. Clique-
width, defined in [4], is the older of the two notions. In several aspects, the theory
of graphs of bounded clique-width is similar to the one of bounded tree-width.
Indeed, bounded tree-width implies bounded clique-width. However, unlike tree-
width, graphs with bounded clique-width include arbitrarily large cliques and
other dense graphs. On the other hand, clique-width is not closed under taking
subgraphs (or minors), just induced subgraphs.

The tree-depth of a graph has been defined in [16], and is equivalent or
similar to notions such as the *vertex ranking number* and the minimum height of
an *elimination tree* [1, 5, 20], etc. Graphs with bounded tree-depth are sparse, and
enjoy strong "finiteness" properties (finiteness of cores, existence of non-trivial

automorphism if the graph is large, well quasi-ordering by subgraph inclusion order). They received almost immediate attention and play a central role in the theory of graph classes of bounded expansion [17].

Graphs of bounded parameters such as clique-width allow us to efficiently solve various optimization problems which are difficult (e.g. NP-hard) in general [3, 6, 12, 11]. However, instead of solving each problem separately, we may ask for results which give a solution to a whole class of problems. We call such results *algorithmic metatheorems*. One of the most famous results of this kind is Courcelle's theorem [2], which states that every graph property expressible in MSO_2 logic of graphs can be solved in linear time on graphs of bounded tree-width. More precisely, the MSO_2 model-checking problem for a graph G of tree-width $tw(G)$ and a formula ϕ, i.e. the question whether $G \models \phi$, can be solved in time $\mathcal{O}(|G| \cdot f(\phi, tw(G)))$. (In the world of parameterized complexity we say that such problems, solvable in time $\mathcal{O}(n^p \cdot f(k))$ for some constant p and a computable function f, where k is some parameter of the input and n the size of the input, are *fixed-parameter tractable (FPT)*.) For clique-width a result similar to Courcelle's theorem holds: MSO_1 model checking is FPT on graphs of bounded clique-width [3].

However, an issue with these results is that, as showed by Frick and Grohe [7] for MSO model checking of the class of all trees, the function f of Courcelle's algorithm is, unavoidably, non-elementary in the parameter ϕ (unless P=NP). This brings the following question: Are there any interesting graph classes where the dependency on the formula is better? Only recently, in 2010, Lampis [15] gave an FPT algorithm for MSO_2 model checking on graphs of bounded vertex cover with elementary (doubly-exponential) dependence on the formula. A current result of Gajarský and Hliněný [9] shows that there exists an FPT algorithm for MSO_2 model checking for graphs of bounded tree-depth, again with elementary dependency on the formula.

Our results. Motivated by the success of tree-depth, we would like to formalize a parameter which extends tree-depth towards a logic-flavoured graph description such as that of clique-width. We start by introducing two such parameters: *shrub-depth* and *SC-depth*. Both of these parameters are based on the notion of *tree-model*, which can be seen as a minimalistic analogue of graph interpretation into a tree. Shrub-depth and SC-depth are then defined in terms of the number of layers (the depth) such a tree-model must have to be able to interpret a given graph.

The first main result of this paper is that the classes of the graphs resulting from an MSO_1 graph interpretation in the class of all finite rooted trees of height $\leq d$, with vertices labelled by a finite set of labels, are exactly the classes of graphs of shrub-depth at most d. This result, in combination with [9], leads to an FPT algorithm for MSO_1 model checking for graphs of bounded shrub-depth (SC-depth) with an elementary dependence on the formula.

Continuing in the same direction we also introduce the notion of m-partite cographs, which are a natural extension of ordinary cographs. (Recall that a

cograph is a graph that can be generated from K_1 by complementations and disjoint unions.) We argue that m-partite cographs represent a smooth intermediate transition from the shrub- and SC-depth to the significantly wider and established notions of clique-width [4] and NLC-width [21]. Indeed, we show that all graphs in any class of shrub-depth d are m-partite cographs with a representation of depth $\leq d$, for suitable m. On the other hand, every m-partite cograph has clique-width at most $2m$.

The second main result of this paper is that the class of m-partite cographs is well-quasi-ordered by the relation of "is an induced subgraph of". This is a significant result, which implies that a) testing whether a graph is an m-partite cograph is an FPT problem, and b) deciding any hereditary property (i.e. property closed under taking induced subgraphs) on this class is an FPT problem.

Paper organization. In Section 2 we give the necessary definitions, including $\mathrm{MSO}_1/\mathrm{MSO}_2$ logics and FO/MSO graph interpretation, which is a specialized instance of the concept of interpretability of logic theories. Section 3 then introduces tree models and, through them, the new invariants shrub-depth and the related SC-depth. Section 4 deals with MSO model checking and interpretability. In Section 5 we investigate the more general concept of m-partite cographs and relate it to other invariants. We conclude with Section 6.

2   Definitions

We assume the reader is familiar with standard notation of graph theory. In particular, all our graphs (both directed and undirected) are finite and simple (i.e. without loops or multiple edges). For a graph $G = (V, E)$ we use $V(G)$ to denote its vertex set and $E(G)$ the set of its edges. We will often use *labelled graphs*, where each vertex is assigned one of some fixed finite set of labels. A forest F is a graph without cycles, and a tree T is a forest with a single connected component. We will consider mainly *rooted forests (trees)*, in which every connected component has a designated vertex called the *root*. The *height* of a vertex x in a rooted forest F is the length of a path from the root (of the component of F to which x belongs) to x and is noted height(x, F). The *height*[5] of the rooted forest F is the maximum height of the vertices of F. Let x, y be vertices of F. The vertex x is an *ancestor* of y, and y is a *descendant* of x, in F if x belongs to the path of F linking y to the corresponding root. If x is an ancestor of y and $xy \in E(T)$, then x is called a *parent* of y, and y is a *child* of x.

_____

[5] There is a conflict in the literature about whether the height of a rooted tree should be measured by the "root-to-leaves distance" or by the "number of levels" (a difference of 1 on finite trees). We adopt the convention that the height of a single-node tree is 0 (i.e., the former view).

Tree-depth. The *closure* $\mathrm{Clos}(F)$ of a forest F is the graph obtained from F by making every vertex adjacent to all of its ancestors. The *tree-depth* $\mathrm{td}(G)$ of a graph G is one more than the minimum height of a rooted forest F such that $G \subseteq \mathrm{Clos}(F)$ [16]. For a proof of the following proposition, as well as for a more extensive study of tree-depth, we refer the reader to [18].

Proposition 2.1. *Let G and H be graphs. Then the following is true:*

- *If H is a minor of G then $\mathrm{td}(H) \leq \mathrm{td}(G)$.*
- *If L is the length of a longest path in G then $\lceil \log_2(L+2) \rceil \leq \mathrm{td}(G) \leq L+1$.*
- *If $\mathrm{tw}(G)$ and $\mathrm{pw}(G)$ denote in order the tree-width and path-width of a graph, then $\mathrm{tw}(G) \leq \mathrm{pw}(G) \leq \mathrm{td}(G) - 1$.*

Clique-width. A *k-expression* is an algebraic expression with the following four operations on vertex-labelled graphs using k labels: create a new vertex with label i; take the disjoint union of two labelled graphs; add all edges between vertices of label i and label j; and relabel all vertices with label i to have label j. The *clique-width* [4] of a graph G equals the minimum k such that (some labelling of) G is the value of a k-expression.

MSO logic and interpretation. We now briefly introduce the *monadic second order logic (*MSO*)* over graphs and the concept of FO (MSO) graph interpretation. MSO is the extension of first-order logic (FO) by quantification over sets:

Definition 2.2 (MSO_1 **logic of graphs**)**.** *The language of MSO_1 consists of expressions built from the following elements:*

- *variables $x, y, \ldots$ for vertices, and X, Y for sets of vertices,*
- *the predicates $x \in X$ and $\mathrm{edge}(x, y)$ with the standard meaning,*
- *equality for variables, quantifiers $\forall, \exists$ ranging over vertices and vertex sets, and the standard Boolean connectives.*

MSO_1 logic can be used to express many interesting graph properties, such as 3-colourability. We also mention MSO_2 logic, which additionally includes quantification over edge sets and can express properties which are not MSO_1 definable (e.g. Hamiltonicity). The large expressive power of both MSO_1 and MSO_2 makes them a very popular choice when formulating algorithmic metatheorems (e.g., for graphs of bounded clique-width or tree-width, respectively).

A useful tool when solving the model checking problem on a class of structures is the ability to "efficiently translate" an instance of the problem to a different class of structures, for which we already have an efficient model checking algorithm. To this end we introduce simple $\mathrm{FO}/\mathrm{MSO}_1$ graph interpretation, which is an instance of the general concept of interpretability of logic theories [19] restricted to simple graphs with vertices represented by singletons.

Definition 2.3. *A* FO (MSO_1) graph interpretation *is a pair $I = (\nu, \mu)$ of FO (MSO_1) formulae (with 1 and 2 free variables respectively) in the language of*

4

graphs, where μ is symmetric (i.e. $G \models \mu(x,y) \leftrightarrow \mu(y,x)$ in every graph G). To each graph G it associates a graph $I(G)$ (by standard abuse of notation), which is defined as follows:

– *The vertex set of $I(G)$ is the set of all vertices v of G such that $G \models \nu(v)$;*
– *The edge set of $I(G)$ is the set of all the pairs $\{u,v\}$ of vertices of G such that $G \models \nu(u) \wedge \nu(v) \wedge \mu(u,v)$.*

This definition naturally extends to the case of vertex-labelled graphs (using a finite set of labels, sometimes called colours) by introducing finitely many unary relations in the language to encode the labelling.

3 Capturing height of graphs

To motivate the definition of shrub-depth, we recall the simple *neighbourhood diversity* parameter introduced by Lampis [15] in his search for graph classes having a faster MSO_1 model checking algorithm: Two vertices u, v are *twins* in a graph G if $N_G(u) \setminus \{v\} = N_G(v) \setminus \{u\}$. The *neighbourhood diversity* of a graph G is the smallest m such that $V(G)$ can be partitioned into m sets such that in each part the vertices are pairwise twins (each part is then either a clique or independent). This basically means that $V(G)$ can be coloured by m labels such that the existence of an edge uv depends solely on the labels of u and v.

Inspired by some subsequent generalizations of neighbourhood diversity, e.g, in [10, 8], our idea is to enrich it with a bounded number of "layers". That is, we bring the following definition, which can also be viewed as a very simplified (or minimalistic) analogue of a graph interpretation (Def. 2.3) into a tree of bounded height:

Definition 3.1 (Tree-model). *We say that a graph G has a* tree-model of m labels and depth d *if there exists a rooted tree T (of height d) such that*

   i. the set of leaves of T is exactly $V(G)$,
  ii. the length of each root-to-leaf path in T is exactly d,
 iii. each leaf of T is assigned one of m labels (T is m-labelled),
 iv. and the existence of a G-edge between $u, v \in V(G)$ depends solely on the labels of u, v and the distance between u, v in T.

The class of all graphs having a tree-model of m labels and depth d is denoted by $\mathcal{TM}_m(d)$.

Note that there is no explicit computability assumption in Definition 3.1.iv; it is implicit from the fact that a tree-model has fixed height and uses a bounded number of labels.

For instance, $K_n \in \mathcal{TM}_1(1)$ or $K_{n,n} \in \mathcal{TM}_2(1)$. Definition 3.1 is further illustrated in Figure 1. It is easy to see that each class $\mathcal{TM}_m(d)$ is closed under complements and induced subgraphs, but neither under disjoint unions, nor under subgraphs. The depth of a tree model generalizes tree-depth of a graph as follows (while the other direction is obviously unbounded, e.g., for cliques):
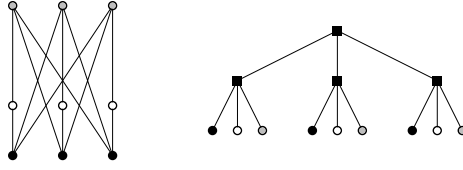
Fig. 1. The graph obtained from $K_{3,3}$ by subdividing a matching belongs to $\mathcal{TM}_3(2)$.

Proposition 3.2. *If G is of tree-depth d, then $G \in \mathcal{TM}_{2^d}(d)$, and $G \in \mathcal{TM}_{2^d}(d-1)$ if, moreover, G is connected.*

In the technical definition of a tree-model, the depth parameter d is much more important (for our purposes, e.g. efficient MSO property testing) than the number of labels m. With this in mind, it is useful to work with a more streamlined notion which only requires a single parameter, and to this end we introduce the following:

Definition 3.3 (Shrub-depth). *A class of graphs $\mathcal{G}$ has shrub-depth d if there exists m such that $\mathcal{G} \subseteq \mathcal{TM}_m(d)$, while for all natural m it is $\mathcal{G} \not\subseteq \mathcal{TM}_m(d-1)$.*

Note that Definition 3.3 is asymptotic as it makes sense only for infinite graph classes; the shrub-depth of a single finite graph is always at most one (0 for empty or one-vertex graphs). Furthermore, it makes no sense to say "the class of all graphs of shrub-depth d".

For instance, the class of all cliques has shrub-depth 1. For more relations of shrub-depth to other established concepts such as cographs or clique-width we refer the reader to Section 5. It is, however, immediate from Definition 3.1 that all graphs in $\mathcal{TM}_m(d)$ have clique-width $\leq m$, and our bounded shrub-depth indeed "lies between" bounded tree-depth and bounded clique-width:

Proposition 3.4. *Let $\mathcal{G}$ be a graph class and d an integer. Then:*

a) If $\mathcal{G}$ is of tree-depth $\leq d$, then $\mathcal{G}$ is of shrub-depth $\leq d$ (cf. Proposition 3.2).
b) If $\mathcal{G}$ is of bounded shrub-depth, then $\mathcal{G}$ is of bounded clique-width.

The converse statements are not true in general.

SC-depth. One can come with yet another, very simple and single-parameter based, definition of a depth-like parameter which is asymptotically equivalent to shrub-depth: Let G be a graph and let $X \subseteq V(G)$. We denote by $\overline{G}^X$ the graph G' with vertex set $V(G)$ where $x \neq y$ are adjacent in G' if (i) either $\{x,y\} \in E(G)$ and $\{x,y\} \not\subseteq X$, or (ii) $\{x,y\} \notin E(G)$ and $\{x,y\} \subseteq X$. In other words, $\overline{G}^X$ is the graph obtained from G by complementing the edges on X.

Definition 3.5 (SC-depth[6]). *We define inductively the class $\mathcal{SC}(n)$ as follows:*

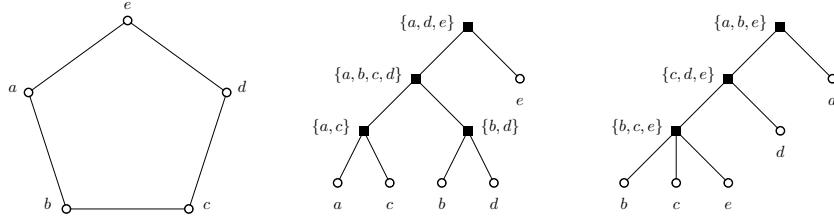---

[6] As the "Subset-Complementation" depth.

Fig. 2. A graph G and two possible SC-depth representations by depicted trees.

- *We let $\mathcal{SC}(0) = \{K_1\}$;*
- *if $G_1, \ldots, G_p \in \mathcal{SC}(n)$ and $H = G_1 \dot{\cup} \ldots \dot{\cup} G_p$ denotes the disjoint union of the G_i, then for every subset X of vertices of H we have $\overline{H}^X \in \mathcal{SC}(n+1)$.*

The SC-depth *of G is the minimum integer n such that $G \in \mathcal{SC}(n)$.*

The SC-depth of a graph G is thus the minimum height of a rooted tree Y, such that the leaves of Y form the vertex set of G, and each internal node v is assigned a subset X of the descendant leaves of v. Then the graph corresponding to v in Y is the complement on X of the disjoint union of the graphs corresponding to the children of v (see Fig. 2).

Theorem 3.6. *Let $\mathcal{G}$ be a class of graphs. Then the following are equivalent:*

- *There exist integers d, m such that $\mathcal{G} \subseteq \mathcal{TM}_m(d)$ (i.e. $\mathcal{G}$ has bounded shrub-depth).*
- *There exists an integer k such that $\mathcal{G} \subseteq \mathcal{SC}(k)$ (i.e. $\mathcal{G}$ has bounded SC-depth).*

The reason we introduce both asymptotically equivalent SC-depth and shrub-depth measures here is that each one brings a unique perspective on the class of graphs we are interested in (and for a yet another, more general, perspective we refer to Section 5).

4 MSO interpretation and model checking

In this section we present the first main result of the paper, Theorem 4.1, which shows that our tree-model (Def. 3.1) indeed fully captures the power of an MSO_1 graph interpretation. While such a result may be expected for FO logic, we believe it is rather surprising in the full scope of MSO logic. The assumption of bounded depth of the target tree is absolutely essential here.

Theorem 4.1. *A class $\mathcal{G}$ of graphs has an MSO_1 graph interpretation in the class of all finite rooted trees of height $\leq d$, with vertices labelled by a finite set of labels, if and only if $\mathcal{G}$ has shrub-depth at most d.*

While the proof of Theorem 4.1 is rather involved by itself, it strongly relies also on the following recent result which is of independent interest:

Proposition 4.2 (Gajarský and Hliněný [9]). *Let T be a rooted tree with each vertex assigned one of a finite set of m labels, and let ϕ be any MSO_1 sentence with q quantifiers. There exists a function[7] $R(q,m,d) \leq exp^{(d)}\big((q + m)^{\mathcal{O}(1)}\big)$ such that the following holds:*

Assume a node $u \in V(T)$ such that the subtree $T_u \subseteq T$ rooted at u is of height d, and denote by $U_1, U_2, \ldots, U_k$ the connected components of $T_u - u$. If there is $I \subseteq \{2, \ldots, k\}$, $|I| \geq R(q,m,d)$, such that there exist label-preserving isomorphisms from U_1 to each U_i, $i \in I$, let $T' = T - V(U_1)$. Then, $T \models \phi \iff T' \models \phi$.

Theorem 4.1, combined with another recent result of [9] — stated here as Theorem 4.3, gives us an MSO_1 model checking algorithm which is in FPT and has an elementary dependence on the size of the formula ϕ. Note that this claim is in contrast with the well known algorithms of Courcelle [2] for MSO_2 and Courcelle et al. [3] for MSO_1 model checking problems on the graphs of bounded tree- and clique-width, both of which were shown to have a non-elementary lower bound in ϕ by Frick and Grohe [7].

Theorem 4.3 (Gajarský and Hliněný [9]). *Assume $d \geq 1$ is a fixed integer. Let T be a rooted tree of height d with vertices labelled by a finite set of m labels, and let ϕ be any MSO_1 sentence with q quantifiers. Then the MSO_1 model checking problem $T \models \phi$ can be solved by an FPT algorithm, concretely in time $\mathcal{O}\big(|V(T)|\big) + exp^{(d+1)}\big((q + m)^{\mathcal{O}(1)}\big)$, which is elementary in the parameters ϕ and m.*

Corollary 4.4. *Assume $d \geq 1$ is a fixed integer. Let $\mathcal{G}$ be any graph class of shrub-depth (or SC-depth) $\leq d$. Then the MSO_1 model checking problem on $\mathcal{G}$, i.e., testing $G \models \phi$ for the input $G \in \mathcal{G}$ and MSO_1 sentence ϕ, can be solved by an FPT algorithm, the runtime of which has an elementary dependence on the parameter ϕ. This assumes G is given on the input alongside with its tree-model of depth d.*

Corollary 4.4 thus nicely complements Theorem 4.3 and its straightforward consequence in fast MSO_2 model checking of graphs of bounded tree-depth. The converse direction of Theorem 4.1 moreover shows that bounded shrub-depth exactly characterizes the largest extent to which faster MSO_1 model checking can be obtained by the means of Theorem 4.3 – the primary motivation of our research here.

Note that the result of Corollary 4.4 assumes we are given a tree-model of G of depth d on the input. It is therefore natural to ask what is the complexity of obtaining such a model. So far, we have not reached much progress in this direction—while we can test in FPT whether a graph belongs to $\mathcal{TM}_m(d)$ (Section 5), we are not yet able to construct a corresponding tree-model (or, alternatively, an SC-depth tree).

---

[7] Here $exp^{(d)}$ stands for the iterated ("tower of height d") exponential, i.e., $exp^{(1)}(x) = 2^x$ and $exp^{(i+1)}(x) = 2^{exp^{(i)}(x)}$.

5 On m-partite cographs

A *cograph*, or *complement-reducible* graph, is a graph that can be generated from K_1 by complementations and disjoint unions. Cographs were introduced independently by several authors in the seventies, and they are exactly those graphs excluding an induced path of length three. The tree representation of a cograph G is a rooted tree T (called *cotree*), whose leaves are the vertices of G and whose internal nodes represent complemented union.

Some generalizations of cographs have been proposed; e.g., bi-cographs [13] or k-cographs [14]. The following generalization we present here is very natural:

Definition 5.1 (m-partite cograph). *An m-partite cograph is a graph that admits an m-partite cotree representation, that is a rooted tree T such that*

- *the leaves of T are the vertices of G, and are coloured by a label from $\{1, \ldots, m\}$,*
- *the internal nodes v of T are assigned symmetric functions $f_v : \{1, \ldots, m\} \times \{1, \ldots, m\} \to \{0, 1\}$ with the property that two vertices x and y of G with respective colours i and j are adjacent iff their least common ancestor v in T has $f_v(i, j) = 1$.*

By extension, the depth *of an m-partite cograph G is the minimum height of an m-partite cotree representation of G.*
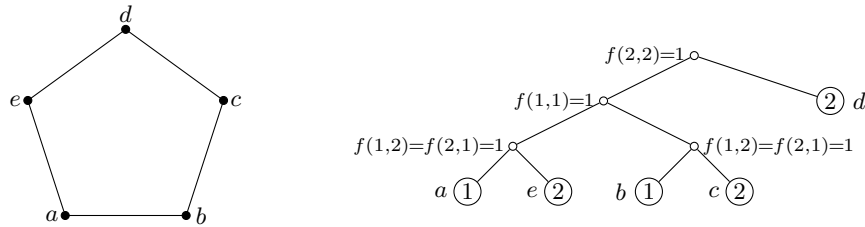

Fig. 3. A 2-partite cotree representation of the graph cycle C_5, with $f(x, y) = 0$ unless otherwise specified.

One can easily deduce from the definition that the graphs in $\mathcal{TM}_m(d)$ are all m-partite cographs of depth $\leq d$. A converse claim is also true (although not immediate).

Theorem 5.2. *Let $\mathcal{G}$ be a graph class. Then the following are equivalent:*

- *There exist integers d, m such that $\mathcal{G}$ only contains m-partite cographs of depth $\leq d$.*
- *The class $\mathcal{G}$ has bounded shrub-depth (or bounded SC-depth, cf. Theorem 3.6).*

It is instructive to look at the general relation of m-partite cographs to shrub-depth and clique-width. The crucial difference between a tree-model and an m-partite cotree representation is in bounding the height of the former. Comparing an m-partite cotree representation to a clique-width expression, roughly saying, the difference lies in the absence of the relabelling operator for the former (this is better seen with the related NLC-width notion [21]). Altogether we get:

Proposition 5.3. *Every m-partite cograph has clique-width at most $2m$.*

Figure 4 summarizes the inclusion hierarchy of the classes we have considered. We give next two examples illustrating the fact that the inclusions indicated in the figure are strict.
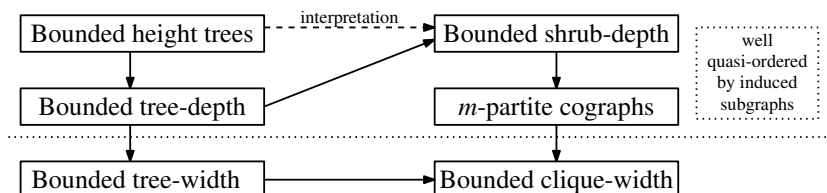

Fig. 4. Hierarchy of graph classes. Arrows mean strict inclusion.

Example 5.4. a) Let H_n denote the graph obtained from the disjoint union of an independent set $\{a_1, \ldots, a_n\}$ and a clique on $\{b_1, \ldots, b_n\}$ by adding all edges $a_i b_j$ such that $i \geq j$. Although each H_n is a 2-partite cograph, the class $\{H_n\}$ has unbounded shrub-depth.

b) The class of all paths has clique-width ≤ 3, while a path of length n is an m-partite cograph if and only if $n < 3(2^m - 1)$.

A *well-quasi-ordering* (or *wqo*) of a set X is a quasi-ordering such that for any infinite sequence of elements $x_1, x_2, \ldots$ of X there exist $i < j$ with $x_i \leq x_j$. In other words, a wqo is a quasi-ordering that does not contain an infinite strictly decreasing sequence or an infinite set of non-comparable elements (i.e. an infinite *antichain*). The existence of a well-quasi-ordering has great algorithmic consequences. For instance, the Robertson-Seymour theorem, which proves that the relation "is a minor of" is a well-quasi-ordering of graphs, implies that for every minor-closed family $\mathcal{C}$ there is a finite set of forbidden minors for $\mathcal{C}$, and hence there is a polynomial time algorithm for testing whether a graph belongs to $\mathcal{C}$. Here we will focus on the quasi-ordering $\subseteq_i$ ("is an induced subgraph of"). A class $\mathcal{C}$ that is closed under taking induced subgraphs is said to be *hereditary*.

Theorem 5.5. *Let m be an integer. The class of m-partite cographs is well-quasi-ordered by the relation $\subseteq_i$ ("is an induced subgraph of").*

10

Corollary 5.6. *a) For every integer m, the class of m-partite cographs is defined by a finite set of excluded induced subgraphs. Hence m-partite cographs are recognizable by an FPT algorithm.*

b) For every hereditary property $\mathcal{P}$ and every integer m, the property $\mathcal{P}$ can be decided by an FPT algorithm on the class of m-partite cographs.

For instance, for every fixed integers k and m, the k-colourability problem can be solved in polynomial time in the class of m-partite cographs.

6   Concluding notes

The main motivation of this paper has been to come up with a notion of "bounded graph depth" which extends the established notion of tree-depth towards dense graphs and parameters similar to clique-width. We have succeeded in this direction with two new, asymptotically equivalent, parameters; shrub-depth and SC-depth. The advantage of the former is that it exactly characterizes the graph classes interpretable in trees of height d, while the latter (SC-depth) outdoes the former with a simpler, single-parameter definition.

Our research topic is also closely related to the class of cographs, and to their natural generalization — m-partite cographs. Saying briefly, graphs of bounded SC-depth are those having m-partite cograph representation of bounded depth. On the other hand, the larger class of all m-partite cographs is "sandwiched" strictly between bounded shrub-depth and bounded clique-width, and it shares several nice finiteness properties with classes of bounded shrub-depth (e.g., well-quasi-ordering under induced subgraphs, which is not true for bounded clique-width classes in general).

The prime algorithmic applications are of two kinds: Firstly, in connection with [9], we obtained an FPT algorithm for MSO_1 model checking for graph classes of bounded shrub-depth which is faster than the algorithm of [3] in the sense that it depends on the checked formula in an elementary way. Secondly, via the well-quasi-ordering property of m-partite cographs, we have proved another algorithmic metatheorem claiming FPT (nonuniform) decidability of all hereditary properties on the classes of m-partite cographs.

Finally, we would like to mention some open questions and directions for future research. Primarily, we do not know yet how to efficiently (in FPT) construct decompositions related to our depth parameters (in this respect our situation is similar to that of clique-width), though we can test existence of such decompositions via Corollary 5.6. We also suggest to investigate the complexity of the isomorphism (or canonical labelling) problem on the classes of bounded shrub-depth, and to try to characterize the maximal graph classes admitting well-quasi-ordering under coloured induced subgraphs. Finally, we remark on the possibility of extending our tools and notions from graphs to general relational structures.

References

1. H. Bodlaender, J. Deogun, K. Jansen, T. Kloks, D. Kratsch, H. Müller, and Z. Tuza. Rankings of graphs. In *WG'95*, volume 903 of *LNCS*, pages 292–304. Springer, 1995.

2. B. Courcelle. The monadic second order logic of graphs I: Recognizable sets of finite graphs. *Inform. and Comput.*, 85:12–75, 1990.

3. B. Courcelle, J. A. Makowsky, and U. Rotics. Linear time solvable optimization problems on graphs of bounded clique-width. *Theory Comput. Syst.*, 33(2):125–150, 2000.

4. B. Courcelle and S. Olariu. Upper bounds to the clique width of graphs. *Discrete Appl. Math.*, 101(1-3):77–114, 2000.

5. J. Deogun, T. Kloks, D. Kratsch, and H. Muller. On vertex ranking for permutation and other graphs. In *STACS'94*, volume 775 of *LNCS*, pages 747–758. Springer, 1994.

6. W. Espelage, F. Gurski, and E. Wanke. How to solve NP-hard graph problems on clique-width bounded graphs in polynomial time. In *WG'01*, volume 2204 of *LNCS*, pages 117–128. Springer, 2001.

7. M. Frick and M. Grohe. The complexity of first-order and monadic second-order logic revisited. *Ann. Pure Appl. Logic*, 130(1-3):3–31, 2004.

8. J. Gajarský. Efficient solvability of graph MSO properties. Master's thesis, Masaryk University, Brno, 2012.

9. J. Gajarský and P. Hliněný. Deciding graph MSO properties: Has it all been told already? Submitted, 2012.

10. R. Ganian. Twin-cover: Beyond vertex cover in parameterized algorithmics. In *IPEC'11*, volume 7112 of *LNCS*, pages 259–271. Springer, 2012.

11. R. Ganian, P. Hliněný, and J. Obdržálek. Clique-width: When hard does not mean impossible. In *STACS'11*, volume 9 of *LIPIcs*, pages 404–415. Dagstuhl Publishing, 2011.

12. M. U. Gerber and D. Kobler. Algorithms for vertex-partitioning problems on graphs with fixed clique-width. *Theoret. Comput. Sci.*, 299(1-3):719–734, 2003.

13. V. Giakoumakis and J.-M. Vanherpe. Bi-complement reducible graphs. *Adv. Appl. Math.*, 18:389–402, 1997.

14. L.-J. Hung and T. Kloks. *k*-cographs are Kruskalian. *Chic. J. Theoret. Comput. Sci.*, 2011:1–11, 2011.

15. M. Lampis. Algorithmic meta-theorems for restrictions of treewidth. In *ESA'10*, volume 6346 of *LNCS*, pages 549–560. Springer, 2010.

16. J. Nešetřil and P. Ossona de Mendez. Tree-depth, subgraph coloring and homomorphism bounds. *European J. Combin.*, 27(6):1024–1041, 2006.

17. J. Nešetřil and P. Ossona de Mendez. Grad and classes with bounded expansion I. Decompositions. *European J. Combin.*, 29(3):760–776, 2008.

18. J. Nešetřil and P. Ossona de Mendez. *Sparsity (Graphs, Structures, and Algorithms)*, volume 28 of *Algorithms and Combinatorics*. Springer, 2012. 465 pages.

19. M. O. Rabin. A simple method for undecidability proofs and some applications. In Y. Bar-Hillel, editor, *Logic, Methodology and Philosophy of Sciences*, volume 1, pages 58–68. North-Holland, Amsterdam, 1964.

20. P. Schaffer. Optimal node ranking of trees in linear time. *Inform. Process. Lett.*, 33:91–96, 1989/90.

21. E. Wanke. *k*-NLC graphs and polynomial algorithms. *Discrete Appl. Math.*, 54:251–266, 1994.

A Appendix

The full concepts and proofs not fitting into the restricted conference paper are presented in this Appendix.

A.1 Shrub-depth: additions to Section 3

Proposition 3.2 *If G is of tree-depth d, then $G \in \mathcal{TM}_{2^d}(d)$, and $G \in \mathcal{TM}_{2^d}(d-1)$ if, moreover, G is connected.*

Proof. Let U be a rooted forest of height $d-1$ such that $G \subseteq \mathrm{Clos}(U)$, and let T be a rooted tree obtained by adding a new root r connected to the former roots of U, and $d' = d$. If G is connected, then U already is a tree, and then we set $T = U$ and $d' = d - 1$.

For $u \in V(T)$ we set a label $c(u) = (j, I)$ such that $dist_T(r, u) = d' - j$ and $I = \{i : \{u, anc_i(u)\} \in E(G)\}$, where $anc_i(u)$ denotes the ancestor of u in T at distance i from u. Notice that $I \subseteq \{1, \ldots, d-1-j\}$ (because of the height of U), and so the total number of distinct $c(u)$ over all $u \in V(U)$ is $2^{d-1} + 2^{d-2} + \cdots + 1 < 2^d$. Let T' be obtained from T as follows: For every node $u \in V(U)$ such that $dist_T(r, u) < d'$, we add to u a new path with the other end denoted by u' such that $dist_{T'}(r, u') = d'$, and set $c(u') = c(u)$.

We claim that this T' with the labels $c(v)$ in the leaves of T' is the desired tree-model of G. Let G' be the graph defined on the leaves of T' as follows; $\{u, v\} \subseteq V(G')$ is an edge of G' iff, for $c(u) = (j_1, I_1)$, $c(v) = (j_2, I_2)$ and $j_1 < j_2$, it holds $dist_{T'}(u, v) = 2j_2$ and $j_2 - j_1 \in I_1$. Then clearly $G' \simeq G$. ☐

Proposition 3.4 *Let $\mathcal{G}$ be a graph class and d an integer. Then:*

a) If $\mathcal{G}$ is of tree-depth $\leq d$, then $\mathcal{G}$ is of shrub-depth $\leq d$ (cf. Proposition 3.2).
b) If $\mathcal{G}$ is of bounded shrub-depth, then $\mathcal{G}$ is of bounded clique-width.

The converse statements are not true in general.

Proof. a) This follows from Proposition 3.2, and the converse cannot be true in general because of, e.g., the cliques.

b) This is since one can straightforwardly translate a tree-model with m labels (and an arbitrary depth) into an m-expression. A counterexample for the converse claim is the class of all paths by Lemma A.5. ☐

Theorem 3.6 *Let $\mathcal{G}$ be a class of graphs. Then the following are equivalent:*

- *There exist integers d, m such that $\mathcal{G} \subseteq \mathcal{TM}_m(d)$ (i.e. $\mathcal{G}$ has bounded shrub-depth).*
- *There exists an integer k such that $\mathcal{G} \subseteq \mathcal{SC}(k)$ (i.e. $\mathcal{G}$ has bounded SC-depth).*

Proof. The proof of this theorem will actually use an intermediate notion of m-partite cographs from Section 5. Hence it will follow from Lemma A.4 and Theorem 5.2. ☐

A.2 MSO$_1$-interpretation: additions to Section 4

Theorem 4.1 *A class $\mathcal{G}$ of graphs has an* MSO$_1$ *graph interpretation in the class of all finite rooted trees of height $\leq d$, with vertices labelled by a finite set of labels, if and only if $\mathcal{G}$ has shrub-depth at most d.*

Before proving this important statement, we need the following technical lemma:

Lemma A.1. *Assume X, Y are label-preserving vertex automorphism orbits in a rooted labelled tree T, and $x \in X$, $y \in Y$ are chosen arbitrarily. If z is the least common ancestor of x, y in T, then the pair (x, y) is determined uniquely up to a label-preserving automorphism of T by the pair of distances $dist_T(x, z)$, $dist_T(y, z)$.*

Proof. All isomorphisms in this proof are label-preserving (w.r.t. the node labelling of T). Firstly, we clarify the meaning of the claim; for any $x_1, x_2 \in X$, $y_1, y_2 \in Y$, and their least common ancestors z_1, z_2 in T, such that $dist_T(x_1, z_1) = dist_T(x_2, z_2)$ and $dist_T(y_1, z_1) = dist_T(y_2, z_2)$, there is an automorphism of T taking the pair (x_1, y_1) onto (x_2, y_2). We carry on the proof by induction on $d = dist_T(x_1, z_1) + dist_T(y_1, z_1)$.

The base case of $d = 0$ is trivial (since $x_1 = y_1$ and $x_2 = y_2$). Consider now an induction step to $d+1$ where $dist_T(x_1, z_1) \geq 1$. Let x_1', x_2' be the parent nodes of x_1, x_2, respectively, and let X' denote the set of parent nodes of all the members of X. Then X' is a vertex orbit of T, too. By inductive assumption, there is an automorphism τ of T taking the pair (x_1', y_1) onto (x_2', y_2). If $\tau(x_1) = x_3$, then x_3 is a child of x_2', and the subtree of T rooted at x_3 is isomorphic to that of x_2 by transitivity. Therefore, we may without loss of generality assume $x_3 = x_2$, and the induction step is complete. $\square$

Notice that Lemma A.1 and its proof can be easily extended to arbitrary k-tuples of tree nodes, which are then uniquely determined up to a T-automorphism by the shape of their Steiner trees.

Proof (of Theorem 4.1). Observe that it is clear from Definition 3.1 that for any fixed integers d, m, the class $\mathcal{TM}_{m-1}(d)$ has an MSO interpretation (or even first order) in the class $\mathcal{T}_m(d)$ of all finite rooted trees of height at most d with each node assigned one of m labels. It thus follows easily that so does every $\mathcal{G}$ of shrub-depth d for a suitable choice of m.

Now, suppose that (α, β) are MSO formulae defining the interpretation of $\mathcal{G}$ in $\mathcal{T}_m(d)$ for suitable m, i.e., every $G \in \mathcal{G}$ is interpreted in some $T_G \in \mathcal{T}_m(d)$ as follows: $V(G) = \{x \in V(T_G) : T_G \models \alpha(x)\}$ and $E(G) = \{xy : x, y \in V(G) \land T_G \models \beta(x, y)\}$. We will construct a tree-model of G of depth d using a bounded number of (new) labels. For technical reasons, we transform β into a closed sentence, $\beta' \equiv \exists x, y\big(L(x) \land L(y) \land \beta(x, y)\big)$, where L is a new label (added to nodes of the tree). We add the label L to precisely two nodes of T_G for which we will need to test adjacency in G.

14

Let $G \in \mathcal{G}$ be a fixed graph and let $T = T_G$, as above. For $u \in V(T)$, denote by $T_u \subseteq T$ the subtree rooted at u. Let q be the number of quantifiers in β', and for $R_i = R(q, m, i)$ from Proposition 4.2, let $R_i' = R_i + 2$. We repeat the following recursively, starting from the leaf nodes of T, going up to the root: For each $u \in V(T)$ such that T_u is of height i, consider the components of $T_u - u$ partitioned into the equivalence classes according to the existence of a label-preserving isomorphism. We prune the number of components in each class to exactly R_i' whenever possible. Let T' be the resulting "reduced" subtree of T. Observe that T' is of bounded size depending only q, m and d, and independent of the size of T.

Suppose x, y is a pair of nodes of T for which we want to test adjacency in G. Let $T[L(x), L(y)]$ denote the tree T in which the new label L has been assigned to precisely $x, y \in V(T)$. Correspondingly, we denote by $T'[L(x), L(y)]$ the reduced tree as described above. Note that up to symmetry, we may always assume $x, y \in V(T')$. Furthermore, while forming $T'[L(x), L(y)]$, we only remove components from label preserving isomorphism classes of $T_u[L(x), L(y)] - u$ of size $> R_i' - 2 = R_i$, thereby accounting for the possibility that some, at most two, components of $T_u[L(x), L(y)] - u$ receive the label L. Then, it follows from Proposition 4.2 that $T[L(x), L(y)] \models \beta' \iff T'[L(x), L(y)] \models \beta'$. Consequently, for each pair x, y, one can determine whether or not it forms an edge in G simply by testing if T' with a suitable assignment of L satisfies β'.

We now describe how to obtain a tree model for G with height d and a bounded number of labels. Starting with the rooted Steiner tree of $V(G)$ in T_G, we construct U by possibly "pushing" the vertices $V(G)$ to leaves at distance exactly d from the root, in order to fulfil Definition 3.1. For each $u \in V(G)$, let $d(u)$ denote the distance it is pushed while forming U and $Or(u)$ be the automorphism orbit of u in T'. We assign the label $(T', Or(u), d(u))$ to each $u \in V(G)$. For any pair $u, v \in V(G)$, $|d(u) - d(v)|$ yields the height of their common ancestor, z, thereby also determining $dist_T(x, z)$ and $dist_T(y, z)$. Subsequently, Lemma A.1 implies that the labels $(T', Or(u), d(u))$ and $(T', Or(v), d(v))$ determine the exact mutual position of u, v within T'. Hence, this tree model can determine the value of $T'[L(x), L(y)] \models \beta'$, which in turn determines $T[L(x), L(y)] \models \beta'$, and consequently, the logical value of $T \models \beta(u, v)$. $\qquad\qquad\square$

Now for the second result:

Corollary 4.4 *Assume $d \geq 1$ is a fixed integer. Let $\mathcal{G}$ be any graph class of shrub-depth (or SC-depth) $\leq d$. Then the MSO_1 model checking problem on $\mathcal{G}$, i.e., testing $G \models \phi$ for the input $G \in \mathcal{G}$ and MSO_1 sentence ϕ, can be solved by an FPT algorithm, the runtime of which has an elementary dependence on the parameter ϕ. This assumes G is given on the input alongside with its tree-model of depth d.*

Proof. Recall that that $\mathfrak{T}_m(d)$ denotes the class of all finite rooted trees of height $\leq d$ with each node assigned one of m labels. By (the easy direction of) Theorem 4.1, for suitable finite m and every $G \in \mathcal{G}$ there is labelled $T_G \in \mathfrak{T}_m(d)$ such that G has a simple MSO_1 interpretation in T_G. This tree T_G is straightforwardly

obtained from the assumed tree-model of G, and the interpretation is given by a pair (α, β) of MSO_1 formulae defining $V(G) = \{x \in V(T_G) : T_G \models \alpha(x)\}$ and $E(G) = \{xy : x, y \in V(G) \wedge T_G \models \beta(x, y)\}$.

We now take the input MSO_1 sentence ϕ, and construct ϕ' by replacing every occurrence of '$\exists x(\dots)$' with '$\exists x(\alpha(x) \wedge \dots)$' and of 'edge$(x, y)$' with '$\beta(x, y)$'. Then $G \models \phi$ iff $T_G \models \phi'$. Hence the rest follows from Theorem 4.3. $\qquad\square$

A.3 m-partite cographs: additions to Section 5

Theorem 5.2 *Let $\mathcal{G}$ be a graph class. Then the following properties are equivalent:*

- *there exist integers d, m such that the class $\mathcal{G}$ only contains m-partite cographs of depth at most d;*
- *the class $\mathcal{G}$ has bounded shrub-depth (or SC-depth, cf. Theorem 3.6).*

We deduce the theorem from the following two lemmas relating m-partite cographs to SC-depth of Def. 3.5.

Lemma A.2. *Assume $G \in \mathcal{K}_n$ then G has a 2^n-partite cotree representation of depth n.*

Proof. We prove the statement by induction over n. If $n = 1$ the statement is obviously satisfied. Assume $G \in \mathcal{K}_{n+1}$. Then there exist $G_1, \dots, G_p \in \mathcal{K}_n$ and a subset X of vertices of G such that $G = \overline{\sum G_i}^X$. By induction, each G_i has a 2^n-partite cotree representation of depth n (where the label of a vertex x is denoted by $\mathrm{label}(x)$). Let $X_i = X \cap V(G_i)$. We define a new labelling of the vertices of G_i by $\mathrm{l}(x) = (\mathrm{label}(x), 0)$ if $x \notin X$ and $\mathrm{l}(x) = (\mathrm{label}(x), 1)$ if $x \in X$. By complementing the Boolean functions defining adjacencies between vertices whose label have the form $(i, 1)$, we get that $\overline{G_i}^{X_i}$ has a 2^{n+1}-partite cotree representation of depth n. Moreover, if $p > 1$, by gluing tree representations and defining a new Boolean function f at the root by $f((i, a), (j, b)) = 1$ if $a = b = 1$ we get that G has a 2^{n+1}-cotree representation of depth $n + 1$. $\qquad\square$

Lemma A.3. *Assume G has an m-partite cotree representation of depth d. Then $G \in \mathcal{K}_{3dm^2}$.*

Proof. We prove the statement by induction over d. If $d = 0$ (i.e. $G = K_1$) then the statement is true. Assume G has an m-partite cotree representation of depth $d + 1$. Then there exist a subset $\{G_1, \dots, G_p\}$ of graphs, each having an m-partite cotree representation of depth d, and a symmetric function $f : \{1, \dots, m\} \times \{1, \dots, m\} \to \{0, 1\}$ such that G is obtained from $\sum G_i$ by making adjacent vertices with label α in G_i and vertices with label β in G_j whenever $i \neq j$ and $f(\alpha, \beta) = 1$. By induction, each of the G_i belongs to $\mathcal{K}_{3dm^2}$. For each i, we proceed to a sequence of complementations of subsets of vertices of G_i in such a way that for every $1 \leq a \leq b \leq m$ such that $f(a, b) = 1$ the set of edges between vertices of G_i with label a and b are complemented. This takes at

16

most $3\binom{m}{2}$ iterations and leads to graph G'_i. Then we consider $\sum G'_i$ and again proceed to a sequence of complementations of subsets of vertices of G in such a way that for every $1 \le a \le b \le m$ such that $f(a,b) = 1$ the set of edges between vertices of G_i with label a and b are complemented. Altogether, we proceeded to at most $3m^2$ operations and obtained the graph G hence $G \in \mathcal{K}_{3(d+1)m^2}$. □

Lemma A.4. *Let m, d be arbitrary positive integers.*

- *All graphs in the class $\mathcal{TM}_m(d)$ are m-partite cographs of depth at most d.*
- *For each m there exists m' such that every m-partite cograph of depth d belongs to the class $\mathcal{TM}_{m'}(d)$.*

In other words, a graph class $\mathcal{G}$ is of shrub-depth d iff there exists m such that all members of $\mathcal{G}$ are m-partite cographs of depth at most d.

Proof. The first direction is trivial as every tree-model of m labels is at the same time an m-partite cograph.

As for the other direction, we have to introduce a little trick. For each internal node v of a cotree representation T of a graph G, there is only a bounded number of possible choices of the function f_v (cf. Def. 5.1) for fixed m. Hence, for every leaf x (of T) we can, in addition to the original label, record the functions f_v for all ancestors v of x. Since the cotree representation T is of bounded depth d, this leads to a finite number of m' possible new labels. Lastly, it remains to "prolong" each branch of T to length exactly d. Hence $G \in \mathcal{TM}_{m'}(d)$. □

m-partite cographs and clique-width. That the class of m-partite cographs has bounded clique-width can be deduced from the connection between clique-width and the classes NLC_m introduced by Wanke [21]. Recall that NLC_m consists of all graphs that can be obtained from single vertices with labels in $\{1, \ldots, m\}$ using the two following operations:

- union of two graphs G_1 and G_2, with addition of all edges between vertices of G_1 with label i and vertices of G_2 with label j whenever (i, j) belongs to a given subset S of $\{1, \ldots, m\} \times \{1, \ldots, m\}$, called *signature*;
- relabelling of the vertices according to some mapping from $\{1, \ldots, m\}$ to $\{1, \ldots, m\}$.

The *NLC-width* of a graph is the minimum m such that the graph belongs to NLC_m. It has been proved in [23] that the NLC-width and the clique-width of a graph G are related by

$$\mathrm{NLC\text{-}width}(G) \le \mathrm{clique\text{-}width}(G) \le 2\,\mathrm{NLC\text{-}width}(G). \tag{1}$$

From this we immediately deduce:

Proposition 5.3 *Every m-partite cograph is of clique-width at most $2m$.*

Proof. It is a straightforward induction over the depth of an m-partite cotree-representation that the class of m-partite cographs is included into NLC_m. The proof then follows from inequality (1). □
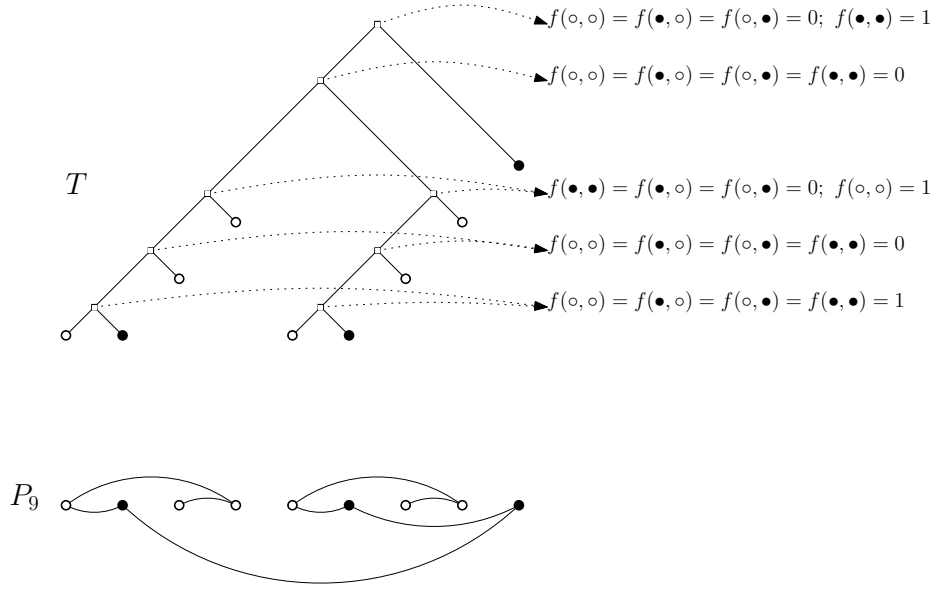
Fig. 5. 2-cotree representation of P_9.

As for **Example 5.4 b)** we have got:

Lemma A.5. *A path on n vertices is an m-partite cograph if and only if $n \leq 3(2^m - 1)$.*

Proof. We first prove by induction over m that if $n = 3(2^m - 1)$ then P_n is an m-partite cograph. The case $m = 1$ is well known. So assume P_n is an m-partite cograph and let T be an m-partite cotree representation of P_n. Take two copies T_1 and T_2 of T, and replace in each of the trees one of the leaves x representing an extremity of the path by a small tree with a root r and two sons: one being x (with the same colour) and the other being a new vertex x' with colour $m + 1$. The function assigned to r being constant and equal to 1. For all the functions assigned to other internal nodes of T_1 and T_2 define $f(i, m+1) = f(m+1, i) = 0$ for every $1 \leq i \leq m + 1$. Let T' be the rooted tree with root r such that r has two sons: a vertex s, with sons T_1 and T_2, and a leaf t, with colour $m + 1$. At the internal vertex s we assign the function f_s such that $f(i, j) = 0$ for every i, j. At the root r we assign the function f_r such that $f(i, j) = 1$ only if $i = j = m + 1$. Then T' is an $(m + 1)$-cotree representation of P_{2n+3} (see Fig 5 for the construction of P_9 as a 2-partite cograph).

We now prove by induction over m that if $n = 3(2^m - 1) + 1$ then P_n is not an m-partite cograph. The case $m = 1$ is again well known. So assume for contradiction that T is an $m+1$-partite cotree representation of $P_{3(2^{m+1}-1)+1} = P_{2n+2}$ and let r be the root of T. As P_{2n+2} is connected, the function f_r assigned to the root is non-zero and there exists i, j (possibly equal) such that some the

18

subtree T_u rooted at a son u of r includes at least a leaf with colour i, the subtree T_v rooted at a son v of r different from u includes a least a leaf with colour j, and $f_r(i, j) = 1$. As none of $K_3, K_{2,2}$, and $K_{1,3}$ are induced subgraphs of P_{2n+2}, we can assume (by symmetry) that T_u includes 1 or 2 leaves with colour i, that T_v includes exactly 1 leaf with colour j. Two cases could occur:

If r has two sons u_1 and u_2 such that T_{u_1} and T_{u_2} includes both one leaf with colour i, $i \neq j$, and only T_v includes a leaf z with colour j. Thus G_z is an m-partite cograph. As this graph includes P_n as an induced subgraph, we get a contradiction.

Otherwise, no son x other than u or v is such that T_x includes a leaf with colour i or j. In this case, we can permute the colours i and j in T_u (and consequently change the functions at interior vertices of T) to ensure $i = j$.

If T_u includes exactly on leaf with colour j then there exists a colour k such that $f_r(j, k) = 1$. It follows that there exists in T exactly one leaf z of colour k, which belongs either to T_u or to T_v. As previously, $G - z$ is an m-partite cograph that includes P_n as an induced subgraph, a contradiction.

Thus T_u includes two leaves with colour j and $f_r(a, b) = 1$ if only if $a = b = j$. By deleting all the vertices of colour j (which form a $K_{1,2}$), we get two connected components, each being a path. The longest one has order at least n. As it is an m-partite cograph, we get a contradiction. $\qquad\square$

Further examples. However, although it is clear that m-partite cographs do not include long induced paths or long induced copaths, the converse is not true: The graph G_n obtained from the path $P_n = (v_1, \ldots, v_n)$ by adding all edges between vertices with even index obviously contains no P_5 and no $\overline{P}_5$. So assume all the G_n are m-partite cographs for some fixed integer m. By assigning to each vertex x with original colour $c(x)$ the colour $(c(x), i(x))$ where $i(x) = 0$ if the index of x is even and $i(x) = 1$ if it is odd, one easily deduce from an m-partite cotree representation of G_n a $2m$-cotree representation of P_n, hence we get a contradiction if we choose $n > 3(2^{2m} - 1)$.

The proof of **Example 5.4 a)** will follow from the next theorem, which shows that every sufficiently large m-partite cograph has many non-trivial automorphisms.

Theorem A.6. *There exists a function $\beta : \mathbb{N}^4 \to \mathbb{N}$ with the following properties: For every integers d, m, k, p, every graph $G \in \mathcal{TM}_m(d)$ of order at least $\beta(d, m, k, p)$ and every colouring $c : V(G) \to \{1, \ldots, k\}$ there exists disjoint subsets $A_1, \ldots, A_p$ of vertices of G such that:*

1. *each of the subsets A_i contains at most $\beta(d, m, k, p)$ vertices,*
2. *for every $1 \leq i < j \leq p$ there exists a c-preserving involutive automorphism $f_{i,j}$ of G globally exchanging A_i and A_j (vertices of $G - A_i - A_j$ being fixed by $f_{i,j}$).*

Proof. Let $G \in \mathcal{TM}_m(d)$ and let Y be a tree-representation of G. To each leaf of Y is associated a vertex of x, with a label label(x) and a colour $c(x)$. Moreover,

to each automorphism of Y preserving both the labels and the colours of the leaves corresponds an automorphism of G. For $n, x, y \in \mathbb{N}$, let $r(n, x, y)$ be the number of unlabelled rooted trees of order at most n with nodes coloured using x colours and tree height y. Let $F(x, y) = x(p-1) + 1$ if $y = 1$ and $F(x, y) = r(F(x, y-1), x, y-1)(p-1) + 1$, otherwise. Assume that Y has at least $F(mk, d)$ leaves. We prove by induction over d that there exists an internal node v of Y such that p isomorphic subtrees of Y are rooted at p sons of v. If $d = 1$ the statement follows from an immediate pigeon-hole argument. Assume that the statement holds for $d = d_0 \geq 1$ and assume $d = d_0 + 1$. Assume for contradiction that no p isomorphic subtrees of Y are rooted at p sons of an internal vertex of Y, and let r be the root of Y. Then, by induction, all the subtrees rooted at a son of r have order at most $F(mk, d-1)$. As there are at most $r(F(mk, d-1), mk, d-1)$ non-isomorphic such subtrees, and has $F(mk, d) > r(F(mk, d-1), mk, d-1)(p-1)$ we conclude that p subtrees rooted at sons of r are isomorphic, a contradiction. □

Such a statement does not hold for m-partite cographs in general, as shown by the construction of Example 5.4 a), which we recall now. For an integer n, consider the graph H_n defined as the graph with vertex set $\{a_i\}_{1 \leq i \leq n} \cup \{b_j\}_{1 \leq j \leq n}$ where a_i is adjacent to b_j if $j \leq i$ and all the b_j form a clique (see Fig 6).
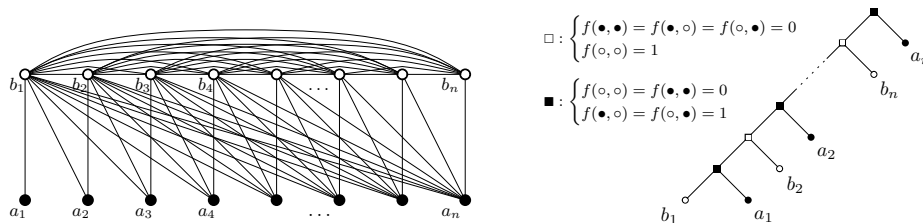

Fig. 6. The graph H_n of Example 5.4 a) is a 2-partite cograph with a single non-trivial automorphism.

Each graph H_n is a 2-partite cograph (how a 2-partite cotree representation of H_n is obtained is shown Fig 6). However, H_n has no non-trivial automorphisms. Indeed, only a_n and b_n have the same degree in H_n hence the only non-trivial automorphism pf H_n exchanges these vertices while fixing all the other vertices. However, according to Theorem A.6 we should find more automorphisms when n is sufficiently large if the class $\{H_n\}$ had bounded shrub-depth.

m-partite cographs and WQO

Theorem 5.5 *Let m be an integer. The class of m-partite cographs is well-quasi-ordered by the relation of "being an induced subgraph of".*

Let T, T' be rooted labeled trees, with labels in a partially ordered set X. A *homeomorphic embedding* $f : T \to T'$ is an injection from $V(T)$ to $V(T')$ such that

- for every x in T it holds label$(x) \leq$ label$(f(x))$;
- for every $x \leq y$ in T it holds $f(x) \leq f(y)$ in T';
- for every x, y in T it holds $f(x \wedge y) = f(x) \wedge f(y)$.

Kruskal proved the following theorem [24]:

Theorem A.7. *Let $\{T_i\}_{i \in I}$ be a family of rooted labelled trees, with labels in a well-quasi-ordered set X. Then there exist $i < j$ such that T_i embeds homeomorphically in T_j.*

Lemma A.8. *Let $T(G)$ and $T(H)$ be m-partite cotree representing m-partite cographs G and H. Assume that $T(H)$ embeds homeomorphically in $T(H)$ in such a way that the root of $T(H)$ is mapped into the root of $T(G)$, that the leaves of $T(H)$ are mapped to leaves of $T(G)$, and that labels are preserved. Then H is (isomorphic to) an induced subgraph of G.*

Proof. Let $f : V(H) \to V(G)$ be the homeomorphic embedding. Each vertex v of $T(G)$ (resp. $T(H)$) is labelled by a symmetric function g_v (resp. h_v) from $\{1, \ldots, m\} \times \{1, \ldots, m\}$ to $\{0, 1\}$. We identify $V(G)$ (resp. $V(H)$) with the set $L(G)$ (resp. $L(H)$) of the leaves of $T(G)$ (resp. $T(H)$). If x and y have respective colours i and j, the same holds for $f(x)$ and $f(y)$. Moreover, as f is injective, $f(x)$ and $f(y)$ are adjacent in G if and only if $g_{f(x) \wedge f(y)}(i, j) = 1$. As $f(x) \wedge f(y) = f(x \wedge y)$ and has $g_{f(x)} = h_x$ we get that $f(x)$ and $f(y)$ are adjacent in G if and only if $h_{x \wedge y} = 1$, that is: if and only if x and y are adjacent in H. It follows that H is isomorphic to the induced subgraph of G induced by $f(L(H))$. □

Proof (of Theorem 5.5). Let $(G_i)_{i \in I}$ be a family of m-partite cographs and let $(T(G_i))_{i \in I}$ be m-partite cotree representations of these graphs. Let X be the set of the pairs (α, β) where $\alpha \in \{0, 1, 2\}$, and where β belongs to $\{1, \ldots, m\}$ if $\alpha = 0$ and to the set of the symmetric functions from $\{1, \ldots, m\} \times \{1, \ldots, m\}$ to $\{0, 1\}$ otherwise. This (finite) set X is trivially partially ordered by $(\alpha, \beta) \leq (\alpha', \beta')$ if $(\alpha, \beta) = (\alpha', \beta')$.

We put labels on the vertices of $T(G)$ as follows: to each leaf with colour i we assign the label $(0, i)$, to each internal vertex v with assigned Boolean symmetric function g_v we assign the label $(1, g_v)$ if v is the root of $T(G)$ and the label $(2, g_v)$, otherwise.

According to Theorem A.7, there exist $i < j$ such that T_i embeds homeomorphically in T_j. It follows, by Lemma A.8 that there exist $i < j$ such that G_i is isomorphic to an induced subgraph of G_j. □

Notice that this result obviously extends if we use a well-quasi-ordered set to label the vertices and consider the natural extension of the induced subgraph relation.

Corollary 5.6 *a)* *For every integer m, the class of m-partite cographs is defined by a finite set of excluded induced subgraphs. Hence m-partite cographs are recognizable by an FPT algorithm.*

b) *For every hereditary property $\mathcal{P}$ and every integer m, the property $\mathcal{P}$ can be decided by an FPT algorithm on the class of m-partite cographs.*

Proof. a) For an integer m, let G be a graph that is not an m-partite cograph, such that every proper induced subgraph of G is an m-partite cograph. Let v be a vertex of G. Consider an m-partite cotree representation T of $G-v$. Let r be the root of T. Relabel every leaf x of T by $2\,\mathrm{label}(x)$ if x is not adjacent to v in G and $2\,\mathrm{label}(x)+1$, otherwise. For each internal node x of T, to which was originally assigned a symmetric function $f_x : \{1,\dots,m\}\times\{1,\dots,m\} \to \{0,1\}$ assign the symmetric function $g_x : \{1,\dots,2m+1\}\times\{1,\dots,2m+1\} \to \{0,1\}$ defined by $g_x(i,j) = f_v(\lfloor i/2\rfloor, \lfloor j/2\rfloor)$ if $i,j < 2m+1$, $g_x(i,j) = 0$ if at least one of i,j is $2m+1$, with exception of $g_r(2i+1,2m+1) = g_r(2m+1,2i+1) = 1$ if $1 \le i < m$. Add to the root r of T a new son with label $2m+1$ for v. Then, the obtained rooted tree together with the functions g_x is a $(2m+1)$-cotree representation of G. It follows that the class $\mathcal{F}_m$ of minimal non m-partite cographs is included in the class of $(2m+1)$-partite cographs. According to Theorem 5.5, the class of $(2m+1)$-partite cographs is well-quasi-ordered by $\subseteq_i$. Hence, as $\mathcal{F}_m$ is an antichain for $\subseteq_i$, it is finite, and checking whether a graph G is an m-partite reduces to testing whether a graph in $\mathcal{F}_m$ is an induced subgraph of G.

b) By Theorem 5.5, there exists a finite set $\mathcal{X}_{\mathcal{P},m}$ of the excluded induced subgraphs for the hereditary property $\mathcal{P}$ over the class of m-partite cographs. This set $\mathcal{X}_{\mathcal{P},m}$ may generally not be computable, but since it is finite, we always get (at least) a nonuniform algorithm. The trick which allows us to test for an induced subgraph in $\mathcal{X}_{\mathcal{P},m}$ by an FPT algorithm, is to use a branch-decomposition of bounded width: Using Proposition 5.3 and [22] we can obtain such a bounded-width decomposition of the input graph G in FPT time, and then use an FPT test for $\mathcal{X}_{\mathcal{P},m}$, e.g., by means of MSO$_1$ [3].

Other References

22. P. Hliněný and S. Oum. Finding branch-decomposition and rank-decomposition. *SIAM J. Comput.*, 38:1012–1032, 2008.
23. Ö. Johansson. Clique-decomposition, NLC-decomposition, and modular decomposition — relationships and results for random graphs. *Congr. Numer.*, 132:39–60, 1998.
24. J. Kruskal. Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. *Trans. Amer. Math. Soc.*, 95(2):210–225, 1960.