Shrub-depth: Capturing Height of Dense Graphs

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

[1] Algorithms and Complexity Group, TU Wien, Austria
rganian@gmail.com
[2] Faculty of Informatics, Masaryk University, Brno, Czech Republic[†]
{hlineny,obdrzalek}@fi.muni.cz
[3] Computer Science Inst. of Charles University (IUUK), Praha, Czech Republic[†,‡]
nesetril@iuuk.mff.cuni.cz
[4] CNRS UMR 8557, École des Hautes Études en Sciences Sociales, Paris, France[‡]
pom@ehess.fr

Abstract. The recent increase of interest in the graph invariant called *tree-depth* and in its applications in algorithms and logic on graphs led to a natural question: is there an analogously useful "depth" notion also for dense graphs (say; one which is stable under graph complementation)? To this end, in a 2012 conference paper, a new notion of *shrub-depth* has been introduced, such that it is related to the established notion of clique-width in a similar way as tree-depth is related to tree-width. Since then shrub-depth has been successfully used in several research papers. Here we provide an in-depth review of the definition and basic properties of shrub-depth, and we focus on its logical aspects which turned out to be most useful. In particular, we use shrub-depth to give a characterization of the lower ω levels of the MSO_1 transduction hierarchy of simple graphs.

ACM subject classification: Theory of computation → Logic, Finite Model Theory; Mathematics of computing → Graph theory.

1 Introduction

In this paper, we are interested in a structural graph parameter that is intermediate between clique-width and tree-depth, sharing the nice properties of both. Clique-width, originating in [6, 8], 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 of bounded clique-width include arbitrarily large cliques and other dense graphs, and the value of clique-width does not change much when complementing the edge set of a graph. Clique-width is not closed under taking subgraphs or minors, only under taking induced subgraphs.

[†] Supported by the research centre Institute for Theoretical Computer Science (CE-ITI); Czech Science foundation project No. P202/12/G061.
[‡] Supported by the project LL1201 CORES of the Ministry of Education of the Czech republic.

As we will see later, clique-width is also closely related to trees and monadic second-order logic of graphs.

The notion of tree-depth of a graph, coined by Nešetřil and Ossona de Mendez [30], is equivalent or similar to some older notions such as the *vertex ranking number* and the minimum height of an *elimination tree* [3, 9, 33], etc. Graphs of small tree-depth are related to trees of small height, and they enjoy strong "finiteness" properties (finiteness of cores, existence of non-trivial automorphism if the graph is large, well-quasi-ordering by subgraph inclusion). The tree-depth notion received almost immediate attention, as it plays a central role in the theory of graph classes of bounded expansion [28, 29]. However, graphs of small tree-depth are necessarily very sparse and the notion behaves badly with respect to, say, graph complementation.

Our search for a structural concept "between clique-width and tree-depth" [19] has originally been inspired by algorithmic considerations: graphs of bounded parameters such as clique-width allow efficient solvability of various problems which are difficult (e.g. NP-hard) in general, e.g. [7, 13, 21, 20]. Highly regarded results in this area are those which, instead of solving one problem, give a solution to a whole class of problems (called *algorithmic metatheorems*). The perhaps most famous result of this kind is Courcelle's theorem [4], 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 $\mathrm{tw}(G) = k$ and a formula ϕ, i.e. the question whether $G \models \phi$, can be solved in time $\mathcal{O}\big(|G| \cdot f(\phi, k)\big)$ (meaning that the problem is *fixed-parameter tractable*, FPT for short). For clique-width a result similar to Courcelle's theorem holds; MSO_1 model checking is FPT on graphs of bounded clique-width [7].

However, an issue with these results is that, as showed by Frick and Grohe [14] 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 interesting graph classes in which the run-time dependency on the formula ϕ is better? For instance, in 2010, Lampis [26] gave an FPT algorithm for MSO_2 model checking on graphs of bounded vertex cover with elementary (doubly-exponential) dependence on the formula. Subsequently, in 2012 Gajarský and Hliněný showed [16] that there exists a linear-time FPT algorithm for MSO_2 model checking of graphs of bounded tree-depth, again with elementary dependence on the formula. Their result is essentially best possible, as shown soon after by Lampis [27]. In order to extend that result towards MSO_1 model checking of (some classes of) dense graphs, one would first need to adjust the clique-width concept towards "bounded depth" (as with tree-depth), which is not a simple task.

The aforementioned paper [16] was not the first one explicitly raising the issue of restricting clique-width towards bounded depth in the literature. In 2012, for example, independently Elberfeld, Grohe and Tantau remarked in a context of expressive power of graph FO logic the following [12]: *One idea is to develop an adjusted notion of clique-width that has the same relation to clique-width* *as tree-depth has to tree-width.* Our concept of shrub-depth [19] has provided a quick positive answer also to the question of [12]. Clique-width-like graph decompositions of limited depth have also been used as a tool by Blumensath and Courcelle in [2] (under the name "$\otimes$-decompositions"). However, some of their technical results which may be interesting in our context have not been published anywhere.

In [19] there have been, in fact, introduced two new structural depth parameters; the *shrub-depth* (Definition 3.3) and the *SC-depth* (Definition 3.5), which are asymptotically equivalent to each other. Since their emergence these have been successfully used in several research papers, and shrub-depth in particular is a subject of ongoing interest in finite model theory of graphs.

For instance, the aforementioned [16] (its full journal version, to be precise) has also extended MSO_2 model checking tractability on graphs of bounded tree-depth to MSO_1 on graph classes of bounded shrub-depth, again with an elementary runtime dependence on the checked formula. Furthermore, [16] has generalized the result of [12] to prove that the expressive power of FO and MSO_1 is the same on classes of bounded shrub-depth.

On another topic, Hliněný, Kwon, Obdržálek and Ordyniak [23] have shown that the tree-depth and shrub-depth concepts of graphs are tightly related to each other via so called vertex-minors. Regarding alternative and generalized views of shrub-depth, DeVos, Kwon and Oum [unpublished] in an ongoing work elaborate on the concept of branch-depth of matroids, and prove that a derived new concept of rank-depth of graphs is asymptotically equivalent to shrub-depth.

Paper organization. Since the core initial paper on shrub-depth [19] has appeared only as a short conference version, we take an opportunity here to give a detailed review of this concept and to provide full proofs of the results of [19] enhanced in light of the current state-of-art. After preliminary definitions of Section 2, this overview of shrub-depth and its structural properties (such as Theorems 3.6, 3.7 and 3.10) constitute Section 3 of this paper.

Subsequent Section 4 focuses on logical aspects of shrub-depth, which have so far been of greatest interest. The main result there (Theorem 4.5) proves that the concept of shrub-depth of a graph class is stable under MSO_1 interpretations and transductions (more precisely, the shrub-depth value does not grow under any non-copying MSO_1 transduction). From that we derive (Theorem 4.8) that the integer values of shrub-depth define the lower ω levels of the MSO_1 transduction hierarchy of simple graphs, which partially answers an open question raised by Blumensath and Courcelle in [1]. We conclude with some remarks and open questions in Section 5.

2 Common Definitions

We assume the reader is familiar with the standard notation of graph theory. In particular, our graphs are finite, undirected 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 write $G \simeq H$ to say that graphs G and H are isomorphic. We will also use *labelled graphs*, where each vertex is assigned one or more of a fixed finite set of labels (in this case, isomorphism implicitly preserves the 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. 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. We also write $y \leqslant x$ in F. 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. The least common ancestor of x and z in F is denoted by $x \wedge z$.

2.1 Width and depth measures

So called width measures play an important role in structural graph theory and in its algorithmic applications. A prototypical width parameter is the *tree-width* of a graph [32] introduced by Robertson and Seymour together with related *path-width*. We refer to [10] for missing definitions and basic properties.

The primary interest of our paper are two other, seemingly unrelated, structural width measures which we define now.

Definition 2.1 (Clique-width [6, 8]). *A* k-expression *is an algebraic expression with the following four operations on vertex-labelled graphs using k labels:*

- *create a new vertex with single label i;*
- *take the disjoint union of two labelled graphs;*
- *add all edges between vertices of label i and label j ($i \neq j$); and*
- *relabel all vertices with label i to label j.*

The clique-width $\mathrm{cw}(G)$ *of a graph G equals the minimum k such that (some labelling of) G is the value of a k-expression.*

Clique-width may be low even on graph classes for which the tree-width is unbounded, such as complete or complete bipartite graphs. Note that Definition 2.1 demands each vertex to carry only one label, while one can allow multiple labels as well. Another possible modification is to allow $i = j$ in the third step. Both these relaxations, while changing values of clique-width for some particular graphs, are nevertheless asymptotically equivalent to the standard clique-width notion of Definition 2.1.

[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).

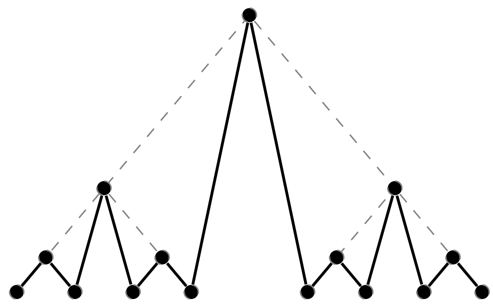

Fig. 1. The path of length n has tree-depth $\log_2(n+2)$, as in the depicted decomposition.

One can, furthermore, define *linear clique-width* (see, e.g., [22]) which has the additional restriction that the union operator is allowed to take only a single vertex as the right-hand operand (i.e., the expression tree is a caterpillar—this is conceptually related to path-width).

A close alternative of clique-width is represented by the NLC classes introduced by Wanke [34]. NLC_m consists of all graphs that can be obtained from single vertices with single labels in $\{1,\ldots,m\}$ using the two following operations:

- disjoint 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 some subset S of $\{1,\ldots,m\} \times \{1,\ldots,m\}$;
- relabelling of the vertices according to some map $\{1,\ldots,m\} \to \{1,\ldots,m\}$.

The *NLC-width* of a graph is the minimum m such that the graph belongs to NLC_m. It has been proved in [25] that the NLC-width and the clique-width of a graph G are related by $\text{NLC-width}(G) \leqslant \mathrm{cw}(G) \leqslant 2\ \text{NLC-width}(G)$.

At last, we briefly mention that another graph measure asymptotically equivalent to clique-width is *rank-width* [31]. Similarly, linear clique-width is asymptotically equivalent to linear rank-width [17].

The second structural measure of our interest is tree-depth.

Definition 2.2 (Tree-depth [30]). *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 $G \subseteq \mathrm{Clos}(F)$.*

Definition 2.2 is illustrated in Figure 1. For a proof of the following proposition, as well as for a more extensive study of tree-depth, we refer the reader to [29].

Proposition 2.3. *Let G and H be graphs. Then the following is true:*

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

2.2 MSO logic on graphs

We now briefly introduce the *monadic second order logic (*MSO*)* over graphs and the concepts of MSO interpretation and transduction. We refer interested readers to, e.g., Courcelle and Engelfriet [5] for further reading. In general, MSO is the extension of first-order logic by quantification over sets. In our paper we deal with the following particular flavour:

Definition 2.4 **(MSO_1 and CMSO_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,*
- *equality for variables, quantifiers $\forall, \exists$ ranging over vertices and vertex sets, and the standard Boolean connectives,*
- *the predicates $x \in X$ and $\mathrm{edge}(x, y)$ with their standard meaning.*

One may also use an arbitrary number of unary predicates on the vertices (as vertex labels). The language of CMSO_1 *(*counting MSO_1*), moreover, adds the predicates* $\bmod_{a,b}$*, such that* $\bmod_{a,b}(X)$ *holds if and only if* $|X| \bmod b = a$*.*

MSO_1 logic can be used to express many interesting graph properties, such as 3-colourability and dominating set as examples. We also briefly mention MSO_2 logic of graphs, which additionally includes quantification over edge sets and can express properties which are not definable in MSO_1 (e.g., Hamiltonicity).

From an algorithmic perspective, MSO logic is particularly useful as the language for describing tractable problems in algorithmic metatheorems (e.g., for aforementioned graphs of bounded clique-width [7] or tree-width [4]). In this respect we consider an $\mathcal{L}$-*model checking* problem in which the input is a graph G, a formula ϕ is the parameter (where ϕ belongs to the considered logic $\mathcal{L}$, such as MSO_1), and the question is whether $G \models \phi$.

A powerful tool, both in theory and in algorithmic metatheorems, is the ability to "efficiently translate" an instance of the model checking problem into another class of instances (for which we, perhaps, already have an efficient model checking algorithm). We start with simple interpretations of undirected graphs.

Definition 2.5. *A* simple MSO_1 graph interpretation *is a pair $I = (\nu, \mu)$ of* MSO_1 *formulae (with* 1 *and* 2 *free variables respectively), such that μ is symmetric (i.e., $G \models \mu(x, y) \leftrightarrow \mu(y, x)$ in every graph G).*[6] *To each graph G it associates a graph $I(G)$ 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 $G \models \nu(u) \wedge \nu(v) \wedge \mu(u, v)$.*

A simple CMSO_1 *graph interpretation is defined analogously.*

[6] We remark that while the question whether μ is symmetric is generally undecidable, we may simply force it to be symmetric, e.g., by using $\mu(x, y) \vee \mu(y, x)$.

For example, a complete graph can be interpreted in any graph (with the same number of vertices) by letting $\nu \equiv \mu \equiv \mathit{true}$, and the complement of a graph has an interpretation using $\mu(x,y) \equiv \neg\,\mathrm{edge}(x,y)$.

Note that, to each CMSO$_1$ formula ϕ, an interpretation $I = (\nu, \mu)$ naturally and efficiently assigns a formula $I(\phi)$ such that $G \models I(\phi) \iff I(G) \models \phi$ holds. Having classes $\mathcal{G}, \mathcal{H}$ of finite graphs, we say that I is a *simple interpretation of* $\mathcal{G}$ *in* $\mathcal{H}$ if the following holds: for every $G \in \mathcal{G}$ there is $H \in \mathcal{H}$ such that $I(H) \simeq G$, and for every $H \in \mathcal{H}$ it holds $I(H) \in \mathcal{G}$.

A more general concept of "logical translation" is that of transductions. Briefly saying, in an addition to a simple interpretation this allows first to add to a graph arbitrary "parameters" (as unary predicates) and to make several disjoint copies of the graph. While a thorough discussion of this concept can be found in [5], here we prefer to give the definition from [1], simplified to target only the graph case:

Definition 2.6 (MSO$_1$ and CMSO$_1$ transduction). *A* basic MSO$_1$ transduction *τ_1 is a triple (χ, ν, μ) such that $(\nu, \mu) = I$ is an MSO$_1$ interpretation, and τ_1 maps a graph G into $I(G)$, or $\tau_1(G)$ is undefined if $G \not\models \chi$.*

The k-copy operation maps a graph G to the graph G^k such that $V(G^k) = V(G) \times \{1, \ldots, k\}$, the subset $V(G) \times \{i\}$ for $i = 1, 2 \ldots, k$ induces a copy of G (there are no edges between distinct copies), and $V(G^k)$ is additionally equipped with a binary relation $\sim$ and unary $P_1, \ldots, P_k$ such that; $(u, i) \sim (v, j)$ for $u, v \in V(G)$ iff $u = v$, and $P_i = \{(v, i) : v \in V(G)\}$.

The p-parameter expansion maps a graph G to the set of all graphs which result by expansion of $V(G)$ by p unary predicates.

Altogether, a many-valued map τ is an MSO$_1$ transduction *(of simple undirected graphs) if it is $\tau = \tau_1 \circ \gamma \circ \varepsilon$ where τ_1 is a basic transduction,[7] γ is a k-copy operation for some k, and ε is the p-parameter expansion for some p. Specially, if $k = 1$ then we call τ a* non-copying transduction.

A CMSO$_1$ *transduction is defined analogously.*

Note, once again, that the result of a transduction τ of one graph is generally a set of graphs, due to involved p-parameter expansion. For a graph class $\mathcal{H}$, the result of a *transduction τ of the class* $\mathcal{H}$ is the union of the particular transduction results, precisely, $\tau(\mathcal{H}) := \bigcup_{G \in \mathcal{H}} \tau(G)$.

3 Capturing Height of Dense Graphs

The concept of tree-depth is commonly used to capture the "height" of other graphs than just trees. Actually, tree-depth can be seen as a bounded-height analogue of tree-width. However, as discussed already in the introduction, the main drawback of tree-depth (as well as of tree-width) is incapability to handle

[7] Here we slightly abuse our simplified terminology since τ_1 can also interpret the predicates $\sim$ and P_i of the k-copy operation. However, the result of τ_1 is always a labelled graph.

dense graphs and some simple graph operations like the complement. Since, on the other hand, clique-width handles dense "uniform" graphs and the complement operation smoothly, it makes good sense to try to modify its definition towards capturing "height" in addition to "width".

Unfortunately, such a direct modification of clique-width seems not possible, and one has to look at other related width measures, namely to the so called neighbourhood diversity and the aforementioned NLC-width for an inspiration.

Before we continue, notice that the requirement to smoothly handle dense graphs and the graph complement operation, naturally means that a new measure cannot be stable under taking non-induced subgraphs.

3.1 Shrub-depth

To motivate the coming definition of shrub-depth, we recall the neighbourhood diversity parameter introduced by Lampis [26] in an algorithmic context: 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 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. This basically means that $V(G)$ can be coloured by m exclusive labels such that the existence of an edge uv depends solely on the colours of u and v.

To stress that the considered labels are exclusive, we shall instead call them *colours*. Inspired by attempts to generalize neighbourhood diversity, e.g, in [18, 15], we come to the idea of enriching the diversity colouring with a bounded number of "layers". This results in the following formalization:

Definition 3.1 (Tree-model). *We say that a graph G has a* tree-model T of m colours 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 colours, and
iv. the existence of a G-edge between $u, v \in V(G)$ depends solely on[8] the colours of u, v and the distance between u, v in T.

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

Point iv. can also be interpreted as that the existence of an edge uv depends on the colours of u, v and the depth of the least common ancestor $u \wedge v$ in T.

For instance, $K_n \in \mathcal{TM}_1(1)$ and $K_{n,n} \in \mathcal{TM}_2(1)$. Definition 3.1 is further illustrated in Figure 2. It is easy to see that each class $\mathcal{TM}_m(d)$ is closed under complements and induced subgraphs (which is our desire), but neither under disjoint unions, nor under subgraphs. Note also that, naturally, one coloured tree T can be a tree-model of several graphs (on the same vertex set), depending on how the edge set is interpreted in point iv.

[8] Notice that a tree-model has fixed height and uses a bounded number of colours, and so there is *no* computability issue involved with the words "depends on" in iv.

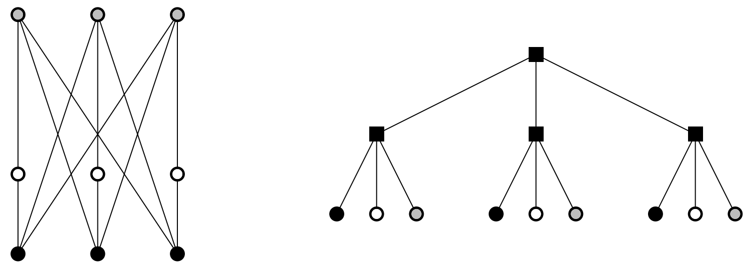

Fig. 2. The graph obtained from $K_{3,3}$ by subdividing a matching belongs to $\mathcal{TM}_3(2)$.

Another interesting observation is the relation of a tree-model to a certain generalization of the NLC classes from Subsection 2.1: imagine that the definition of NLC_m is allowed to make disjoint union of an arbitrary number of graphs, but the depth of the construction tree is bounded by $\leqslant d$. If we, furthermore, forbid the relabelling operation, then the result coincides with the class $\mathcal{TM}_m(d)$. And even if relabellings are allowed in NLC_m, thanks to bounded depth of the construction we can encode all label changes in the leaf colours anyway.

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):

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

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 colour $c(u) = (j, I)$ 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 $I \subseteq \{1, \dots, d-1-j\}$ (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 colours $c(v)$ 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 $G' \simeq G$. $\square$

When dealing with tree-models of graph classes (e.g., in model checking or in transductions), the depth parameter d is asymptotically much more important than the number of colours m. With this in mind, it is useful to work with a more streamlined notion which only requires a single parameter d, 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 $\mathcal{G} \nsubseteq \mathcal{TM}_{m'}(d-1)$.*

In a wider sense, $\mathcal{G}$ is of bounded shrub-depth *if there exist integers d, m such that $\mathcal{G} \subseteq \mathcal{TM}_m(d)$.*

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. On the other hand, it will follow from Theorem 3.7 that the class of all paths has unbounded (infinite) shrub-depth. Now we argue that this new notion is indeed "intermediate" between tree-depth and clique-width (and even linear 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 $\leqslant d$, then $\mathcal{G}$ is of shrub-depth $\leqslant d$.
b) If $\mathcal{G}$ is of bounded shrub-depth, then $\mathcal{G}$ is of bounded linear 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 class of all cliques.

b) We remark that it is trivial to see that $\mathcal{G}$ is of bounded clique-width. Here we even show how to straightforwardly translate a tree-model with m colours and depth d into a linear (caterpillar-shaped) $m(d+1)$-expression: Let $v_1, \ldots, v_n$ be any (usual) left-to-right ordering of the leaves of a tree-model T of some G. The expression is constructed inductively for $i = 1, \ldots, n$ as follows:

- a vertex v_i is created and added with a (currently unique) colour $(c, 0)$ where $c = c(u_i)$ is its colour in T,
- whenever colour c is to be adjacent to colour c' at distance $2d$ in the model T, the expression adds all edges between the colours $(c, 0)$ and (c', d), and
- for $2d'$ being the distance from v_i to v_{i+1} in T, the expression changes all colours (c, d) with $d < d'$ to (c, d').

A counterexample to the converse claim is, e.g., the class of all paths by Theorem 3.7. □

3.2 SC-depth

One can come up 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 $\{x, y\} \nsubseteq X$, 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[9]). *We define inductively the class $\mathcal{SC}(n)$ as follows:*

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

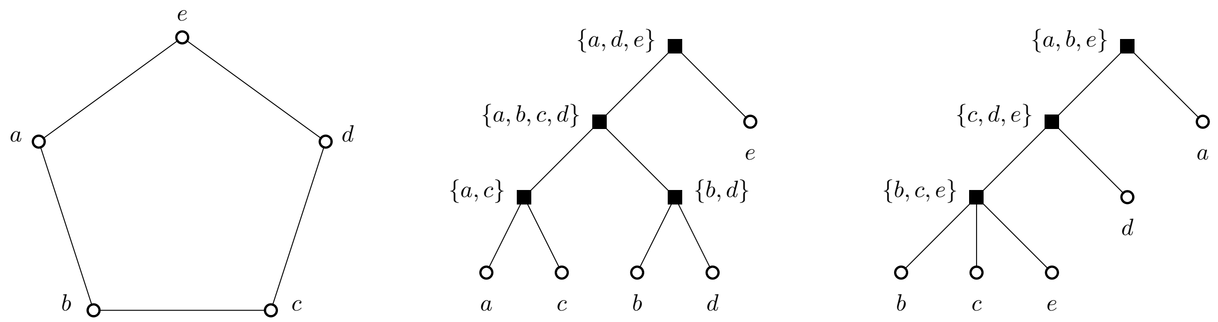

Fig. 3. A graph G and two possible SC-depth representations by depicted trees.

- *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 Figure 3).

The reason we introduce both asymptotically equivalent SC-depth and shrub-depth measures here is that each brings a unique perspective on the classes of graphs we are interested in (see e.g. [23]).

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

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

More precisely, $\mathcal{TM}_m(d) \subseteq \mathcal{SC}(dm(m+1))$ *and* $\mathcal{SC}(k) \subseteq \mathcal{TM}_{2^k}(k)$*.*

Proof. We prove the forward implication by induction on d. In the degenerate base case $d = 0$, it is trivially $\mathcal{TM}_m(0) = \{K_1\} = \mathcal{SC}(0)$. Assume now $G \in \mathcal{TM}_m(d+1)$ for some $d \geqslant 0$. By Definition 3.1, there exist a set $R \subseteq \{1, \ldots, m\}^2$, an integer $c \geqslant 1$ and graphs $G_1, \ldots, G_c \in \mathcal{TM}_m(d)$ (actually subgraphs of G induced by the leaf sets of the root-subtrees in the respective tree-model of G) such that the following holds: G results from the disjoint union $G_1 \cup \cdots \cup G_c$ by adding those edges uv for which u and v belong to distinct graphs among $G_1, \ldots, G_c$, and the pair of colours of u, v (in any order) belongs to R.

By the induction assumption, we have got $G_1, \ldots, G_c \in \mathcal{SC}(k_0)$ for some (computable) integer k_0. For each of these graphs G_ℓ, $\ell \in \{1, \ldots, c\}$, we successively complement edges on the following subsets of vertices:

- for each $(i, i) \in R$, on the set $X^i_\ell \subseteq V(G_\ell)$ of the vertices of G_ℓ of colour i,

– for each $(i,j) \in R$, $i<j$, on the set $X^i_\ell \cup X^j_\ell$ (defined as above), then on the set X^i_ℓ itself and then on X^j_ℓ itself.

Observe that at most $m + 3\binom{m}{2}$ complement operations are applied to each G_ℓ, and this number can be reduced down to $m + \binom{m}{2} = \binom{m+1}{2}$ by skipping possible repeated complements. Denoting by G'_ℓ the graph obtained in this way from G_ℓ hence, by Definition 3.5, we get $G'_1, \dots, G'_c \in \mathcal{SC}(k_1)$ where $k_1 = k_0 + \binom{m+1}{2}$.

Effectively, in each G_ℓ we have complemented the edges whose colour pairs belong to R. In the next step we make the disjoint union $G' := G'_1 \cup \cdots \cup G'_c$ and repeat the same complementation procedure on this global level. Namely:

– for each $(i,i) \in R$, on the set $X^i \subseteq V(G')$ of the vertices of G' of colour i,
– for each $(i,j) \in R$, $i<j$, on the set $X^i \cup X^j$, then on X^i and then on X^j.

Denoting the resulting graph by G'', we similarly get $G'' \in \mathcal{SC}(k_2)$ where $k_2 = k_1 + \binom{m+1}{2} = k_0 + m(m+1)$. It remains to routinely verify that $G'' \simeq G$.

As for the backward implication, we directly construct a tree-model for each graph $G \in \mathcal{SC}(k)$. By Definition 3.5, $G \in \mathcal{SC}(k)$ can be constructed along a rooted tree T such that the leaf set of T is $V(G)$ and each internal node t of T is associated with a complement set X_t (which is a subset of the descendant leaves). We assign the leaf colours as follows. Let $v \in V(G)$ be a leaf of T, and $t_0 = v, t_1, \dots, t_k = r$ be the path from v to the root r of T. We colour v with the binary vector $(a_i)_{i=1}^k$ such that $a_i = 1$ iff $v \in X_{t_i}$.

By Definition 3.5, uv forms an edge of G, if and only if the pair $\{u,v\}$ belongs to an odd number of the complement sets X_t over whole T. This can easily be determined from the colours of u and v and from the depth of their least common ancestor in T. Consequently, $G \in \mathcal{TM}_{2^k}(k)$. □

3.3 Long paths

For graphs of small tree-depth a characteristic property is the absence of long paths as subgraphs, cf. Proposition 2.3 b). This is obviously false for classes of small shrub-depth since those, in particular, include all cliques and bicliques. Although, one can restrict induced paths in every class $\mathcal{TM}_m(d)$, as follows.

Theorem 3.7. *Let $\ell = 3 \cdot 2^m - 4$ and P_ℓ denote the path of length ℓ, i.e., on $\ell+1$ vertices. Then $P_\ell \in \mathcal{TM}_m(2m+1)$, but for any d we have $P_{\ell+1} \notin \mathcal{TM}_m(d)$.*

In particular, there exist no d, m such that $\mathcal{TM}_m(d)$ would contain all paths.

Proof. We start with the construction of $P_\ell \in \mathcal{TM}_m(2m+1)$ that is, of an appropriate tree-model T_m of P_ℓ, by induction on m. We shall maintain a special property that each end of P_ℓ is represented in T_m by a leaf which has no siblings, i.e., its parent is of degree 2. As the base case we use the tree-model T_1 of $m=1$ colours and depth $2m+1=3$ from the left-hand side of Figure 4. (Note that although $P_2 \in \mathcal{TM}_1(2)$, we use an extra level in T_1 to achieve our property.)

We now construct T_{m+1} for $m \geqslant 1$. Let u and v be the ends of P_ℓ, and recall that each of u, v has no siblings in T_m. We create a sibling u_1 of u in T_m and

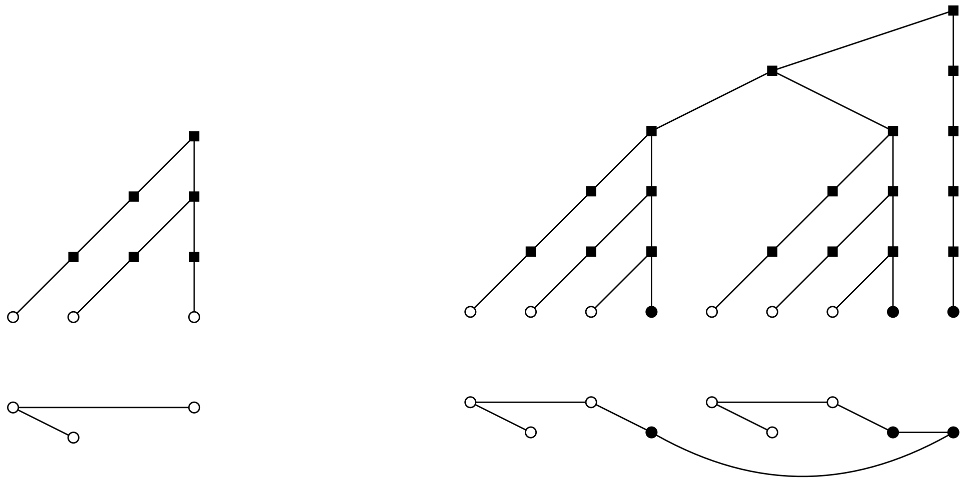

Fig. 4. Tree-models (top) for small paths (below), as used in the proof of Proposition 3.7. Left: with 1 colour and depth 3 for P_2. Right: 2 colours and depth 5 for P_8.

assign u_1 a new colour $m+1$. This intermediate tree-model U_m can represent $P_{\ell+1}$ with the ends u_1, v and, see Figure 4 right, the desired model T_{m+1} follows:

- for U_m and its disjoint copy U'_m, add a common ancestor q of their roots,
- create a rooted path of length $2m+3$, with the root r and the only leaf w of colour $m+1$, and make q another son of r.

Clearly, T_{m+1} is a tree-model of $m+1$ colours and depth $2m+3 = 2(m+1)+1$, and it can represent the edges u_1w and u'_1w but not $u_1u'_1$. Thus T_{m+1} makes a tree-model of $P_{\ell'}$ for $\ell' = 2(\ell+1)+2 = 3 \cdot 2^{m+1} - 4$.

In the converse direction we start with an easy observation for $m = 1$; $P_3 \notin \mathcal{TM}_d(1)$ for any d (this follows from the folklore fact that the path on 4 vertices is not a cograph, too). The proof can then be finished by induction over $m \geqslant 1$, provided that we establish the following contrapositive claim: if $P_{2\ell+5} \in \mathcal{TM}_{m+1}(d)$ for any $\ell, m, d \geqslant 1$, then $P_{\ell+1} \in \mathcal{TM}_m(d)$.

So fix ℓ and m, and assume $G := P_{2\ell+5} \in \mathcal{TM}_{m+1}(d)$ and T is a corresponding tree-model of $m+1$ colours and minimum possible height d. In this proof we denote by T_x the subtree of T rooted at a node x. As d is minimum and $P_{2\ell+5}$ is connected, there exist distinct sons u, v of the root of T and colours i, j (possibly equal), such that T_u includes at least one leaf with colour i and T_v at least one leaf with colour j, and the colour pair (i, j) at distance $2d$ determines an edge.

We let $J \subseteq G$ denote the subgraph formed only by those edges which are determined by the colour pair (i, j) at distance $2d$ in T, i.e., $xy \in E(J)$ iff the colours of x, y are i, j in T and the only common ancestor of x, y is the root of T.

If $i = j$, then we claim that there cannot be two non-incident edges in J. Indeed, this would necessarily mean that J contains $K_{2,2}$, but $K_{2,2} \nsubseteq G$. Hence J is K_2 or $K_{1,2}$ and there exist at most three vertices of colour i altogether, and in either case one subpath in $G - V(J)$ is of length at least $\lceil (2\ell+5-4)/2 \rceil = \ell+1$. Hence $T - V(J)$ gives a tree-model of $P_{\ell+1}$ of m labels.

We now examine the other possibility $i \neq j$. First, we observe that if x_1y_1, x_2y_2 are non-incident edges of J such that x_1, x_2 are of the same colour, then the only common ancestor of x_1, x_2 is the root of T. Otherwise, we would get a contradiction that $K_{2,2} \subseteq J$. Second, we argue that there cannot be three pairwise non-incident edges x_1y_1, x_2y_2, x_3y_3 in J (where x_1, x_2, x_3 are of the same colour). If this happened, then (say) for the vertex y_1 at least two of the vertices x_1, x_2, x_3 would have only one common ancestor with y_1, the root of T. Consequently, y_1 would have at least two neighbours in the set $\{x_1, x_2, x_3, y_1, y_2, y_3\}$, and the same would symmetrically hold for all the members of this set, contradicting the fact that J is acyclic.

Therefore, J is a path of length at most 4, or J consists of two components isomorphic to K_2 or $K_{1,2}$. Moreover, if there exist a leaf z of colour i or j in T which is not incident to an edge of J, then J has no two non-incident edges and all such leaves (of colour i or j) not incident to $E(J)$ are of the same colour, as can be easily checked.

We first consider the case that J has one component. If it is K_2 or $K_{1,2}$ then, by the previous, all the leaves of T coloured i (say) are incident to the one or two edges of J. As above (in the case of $i = j$) we can now argue that $T - V(J)$ gives a tree-model of $P_{\ell+1}$ of m labels. If, on the other hand, J is P_3 or P_4, then all the leaves of T coloured i or j are incident to the edges of J. We form a new tree-model T' by removing from T all the leaves of colours i, j (i.e., incident to the edges of J) and adding arbitrarily one new leaf of colour i. Then T' of m labels models a path $P_{2\ell+1}$ (or $P_{2\ell+2}$).

We are left with the case of J consisting of two components, such that all the leaves of T coloured i or j are incident to the edges of J. If any of the subpaths of $G - V(J)$ is of length at least $\ell + 1$, then we are again done. Otherwise, we can choose one component J_1 of J such that $G - V(J_1)$ contains a subpath G' of length at least $\ell + 3$. We denote by J_2 the other component of J (presumably $J_2 \subseteq G'$), and form a new tree-model T' by restricting T to the leaves from G', removing the leaves of J_2 and adding arbitrarily one leaf of colour i (recall that no vertex of $G' - V(J_2)$ has colour i or j). Hence T' of m labels models a path $P_{\ell+1}$ (or $P_{\ell+2}$). $\square$

The combinatorial result in Theorem 3.7 has interesting relations also to logical questions (see Section 4). For instance, in respect of the research of MSO-orderable graphs by Blumensath and Courcelle [2], note that in the class of all finite paths one can easily define a linear ordering by an MSO_1 formula. Hence it immediately follows from a characterization given in [2] that the class of all finite paths cannot have bounded shrub-depth. The advantage of our Theorem 3.7 (occuring already in [19]) is that it gives exact combinatorial bounds. On the other hand, Theorem 3.7 together with further Theorem 4.1 imply the related result of [2] that infinite graph classes of bounded shrub-depth are not MSO_1-orderable.

Note, however, that graph classes of bounded shrub-depth are not asymptotically related to those excluding long induced subpaths; in the opposite direction the situation here is very different than in Proposition 2.3 b). As an example

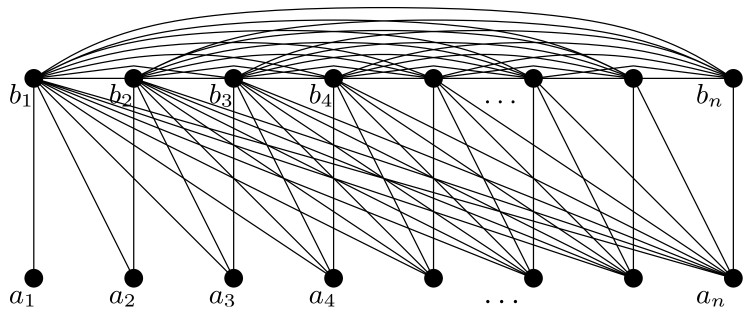

Fig. 5. An example of a graph class not containing any induced subpaths of length 3, which has unbounded shrub-depth. In fact, these graphs even are so called threshold graphs (a special case of small linear clique-width) – view the vertices in the backward order $a_n, b_n, a_{n-1}, b_{n-1}, \dots, a_1, b_1$.

we mention the graph class from Figure 5 which contains no induced subpaths of length 3. One can give a direct combinatorial proof that this class is of unbounded shrub-depth (similarly as for Theorem 3.7), but we skip it here since this fact follows from aforementioned [2] (the graph of Figure 5 is FO-orderable) or, alternatively, from a combination of results of [23].

3.4 Induced subgraphs characterization

Lastly in this section, we provide yet another characterization of the classes defined previously. In a nutshell, we are going to show that each of these classes can be characterized by a finite list of *forbidden induced subgraphs*. A nice consequence of this finding is that membership in each of the classes can be tested in polynomial time. The tool we use here is well-quasi-ordering.

A class or property is said to be *hereditary* if it is closed under taking induced subgraphs. A *well-quasi-ordering* (or *wqo*) of a set X is a quasi-ordering on X such that for any infinite sequence of elements $x_1, x_2, \dots$ of X there exist $i < j$ with $x_i \leqslant x_j$. In other words, a wqo is a quasi-ordering that does not contain an infinite strictly decreasing sequence or an infinite set of incomparable elements. We are going to use the following folklore result:

Theorem 3.8 (Ding [11]). *Let $m \in \mathbb{N}$ be an integer and C be a finite set of colours. The class of the graphs not containing a path on m vertices as a subgraph and with vertices coloured by C is well-quasi-ordered under the colour-preserving induced subgraph order $\subseteq_i$.*

Corollary 3.9. *Let $\mathcal{S}$ be a graph class of bounded shrub-depth, such that their vertices are coloured from a finite set C of colours. Then $\mathcal{S}$ is well-quasi-ordered under the colour-preserving induced subgraph order.*

Proof. Consider an infinite sequence $(G_1, G_2, \dots) \subseteq \mathcal{S}$, and the corresponding tree-models $(T_1, T_2, \dots)$. Let T_i^+, $i = 1, 2, \dots$, denote the rooted tree with leaf labels composed of the colours of T_i and the colours of G_i. By Theorem 3.8,

$T_1^+, T_2^+, \ldots$ of bounded diameter is WQO under rooted coloured subtree relation, and, consequently, so are the coloured graphs $G_1, G_2, \ldots$, as desired. $\square$

The advertised result now follows by a simple twist as follows.

Theorem 3.10. *For every integers d, m, there exists a finite set of graphs $\mathcal{F}_{d,m}$ (the forbidden subgraphs) such that a graph G belongs to $\mathcal{TM}_m(d)$ if and only if G has no induced subgraph isomorphic to a member of $\mathcal{F}_{d,m}$.*

Similarly, for every n there exists a finite set of graphs $\mathcal{F}'_n$ such that $G \in \mathcal{SC}(n)$ if and only if G has no induced subgraph isomorphic to one of $\mathcal{F}'_n$.

Proof. We let $\mathcal{F}_{d,m}$ be the (isomorphism-free) set of graphs H such that $H \notin \mathcal{TM}_m(d)$ but $H - v \in \mathcal{TM}_m(d)$ for every $v \in V(H)$. By this definition, no member of $\mathcal{F}_{d,m}$ is a proper induced subgraph of another member. Hence it is enough to argue that $\mathcal{F}_{d,m}$ is wqo to conclude that $\mathcal{F}_{d,m}$ is finite.

The latter follows from an easy observation: if $H - v \in \mathcal{TM}_m(d)$ for some $v \in V(H)$, then $H \in \mathcal{TM}_{2m+1}(d)$. Indeed, we take a tree-model of $H - v$, add arbitrarily a new leaf of a unique new colour for v and annotate with an extra bit the colours of all leaves which are neighbours of v in H. The result is a tree-model for H with $2m + 1$ colours. Consequently, $\mathcal{F}_{d,m} \subseteq \mathcal{TM}_{2m+1}(d)$ and the wqo property follows from Corollary 3.9.

The second claim is proved analogously. We let $\mathcal{F}'_n$ be the (isomorphism-free) set of graphs H such that $H \notin \mathcal{SC}(n)$ but $H - v \in \mathcal{SC}(n)$ for every $v \in V(H)$. By Theorem 3.6, $\{H - v : H \in \mathcal{F}'_n\} \subseteq \mathcal{TM}_{2^n}(n)$, and so $\mathcal{F}'_n \subseteq \mathcal{TM}_{2^{n+1}+1}(n)$ by the previous paragraph. The wqo property again follows from Corollary 3.9. $\square$

The "obstacle" sets $\mathcal{F}_{d,m}$ and $\mathcal{F}'_n$ of Theorem 3.10 are not only of mathematical interest, but also have algorithmic consequences. Namely, in connection with established algorithms they allow for efficient membership testing of these classes. Note, however, that we do not provide an algorithmic construction of the sets $\mathcal{F}_{d,m}$ and $\mathcal{F}'_n$, and so we only prove an existence of the respective algorithms for each specific values of d, m and n (in parameterized complexity theory this is formally called *nonuniform FPT*).

Corollary 3.11. *The problems to decide, for a given graph G, whether $G \in \mathcal{TM}_m(d)$ and whether $G \in \mathcal{SC}(n)$, are fixed-parameter tractable with respect to the parameters d, m and n, respectively.*

Proof. We provide a proof for the problem of $G \in \mathcal{TM}_m(d)$, while that of $G \in \mathcal{SC}(n)$ is very similar. As mentioned before, the class $\mathcal{TM}_m(d)$ is of bounded clique-width (namely, $2m$ is a trivial upper bound). Therefore, one can use [24] to compute in FPT an approximate expression of G of clique-width depending only on m or to correctly conclude that $G \notin \mathcal{TM}_m(d)$. In the former case, one can then call the algorithm of [7] to test whether any member of $\mathcal{F}_{d,m}$ is an induced subgraph of G. Based on the outcome, the correct decision about $G \in \mathcal{TM}_m(d)$ is easily made. $\square$

4 Shrub-depth and MSO Transductions

While in the previous section we have focused on establishing basic combinatorial properties of shrub-depth and SC-depth, now we shift our attention towards their logical aspects. The final outcome will be the finding that (a slight technical adjustement of) tree-models of depth d precisely capture the d-th finite level of the MSO_1 transduction hierarchy of simple undirected graphs, for all $d \in \mathbb{N}$. For that we start by showing that shrub-depth indeed goes well with simple MSO_1 interpretations.

4.1 Stability under interpretations

We again turn to classical clique-width for an inspiration: graph classes of bounded clique-width have MSO_1 interpretations into the class of all coloured rooted trees and, in turn, graph classes having an MSO_1 interpretation into those of bounded clique-width still have bounded clique-width (although the bound on their clique-width is generally much higher).

In one direction, shrub-depth has been defined using (Definition 3.1) a very special form of a simple MSO_1 interpretation. In the other direction, we can go even further than with clique-width itself (cf. also Section 4.3): the bound on shrub-depth is preserved exactly (and not only asymptotically) under any CMSO_1 interpretations. In other words, the precise height of a tree is absolutely essential for CMSO_1 interpretability. The full formal statement follows.

Theorem 4.1. *A class $\mathcal{G}$ of graphs has a simple CMSO_1 interpretation in a class of finite coloured rooted trees of height at most d, if, and only if, $\mathcal{G}$ has shrub-depth at most d.*

The 'if' direction of Theorem 4.1 follows immediately from Definition 3.1: for any m, the class $\mathcal{TM}_m(d)$ has a simple MSO_1 interpretation (or even FO interpretation) in the class of m-coloured tree-models of depth d. Hence we now give a proof of the 'only if' direction of Theorem 4.1 consisting of the following sequence of three technical claims.

Lemma 4.2 (Gajarský and Hliněný [16]). *There exists a function*[10] $R(q,m,d) \leqslant \exp^{(d)}\big((q+m)^{\mathcal{O}(1)}\big)$ *over the integers such that the following holds.*

Let T be a rooted tree with each vertex assigned one of at most m colours, and let ϕ be any CMSO_1 sentence with q quantifiers, such that the least common multiple of the b values of all $\mathrm{mod}_{a,b}$ predicates in ϕ is at most m. Take any 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 $T_u - u$ (their roots are thus all the k sons of u).

Assume that there exists a (sufficiently large) subset of indices $I \subseteq \{2, \ldots, k\}$, $|I| \geqslant R(q,m,d)$, such that there are colour-preserving isomorphisms from U_1 to

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

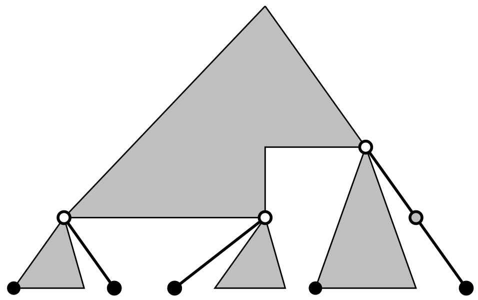

Fig. 6. An illustration of the operation of growing leaves (black dots) from original nodes (white dots) of the depicted rooted tree of height d. All the newly grown leaves have the same distance d from the root, and they may coincide with the original nodes if they already have had distance d from the root.

each U_i, $i \in I$. Then the subtree $T' = T - V(U_1)$ behaves the same with respect to ϕ as T, precisely, $T \models \phi \iff T' \models \phi$.

The operation of obtaining T' from T (*reduction*) as in Lemma 4.2 will be useful in the following setting. Assume we apply this reduction repeatedly in a bottom-up recursion on T. Presicely, let $R' : \mathbb{N} \to \mathbb{N}$. For each $w \in V(T)$ such that T_w is of height i, consider the components of $T_w - w$ partitioned into the equivalence classes according to the existence of a colour-preserving isomorphism. We prune the number of components in each class to exactly $R'(i)$ whenever exceeding this number. Let T'' be the resulting reduced subtree of T. Then we say that T is *R'-reduced to T''*. Observe that T'' is of bounded size depending only on R' and d, and independent of the size of T.

Now we continue with the technical claims leading to Theorem 4.1.

Lemma 4.3. *Let T be a coloured rooted tree. Assume X, Y are colour-preserving vertex automorphism orbits of T, and $x_1, x_2 \in X$ and $y_1, y_2 \in Y$ are chosen arbitrarily. Let $z_i = x_i \wedge y_i$, $i = 1, 2$, denote the least common ancestor of x_i, y_i in T. If $dist_T(x_1, z_1) = dist_T(x_2, z_2)$ and $dist_T(y_1, z_1) = dist_T(y_2, z_2)$, then there is a colour-preserving automorphism of T taking the pair (x_1, y_1) onto (x_2, y_2).*

Proof. For start, all isomorphisms in this proof are colour-preserving . 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 from d to $d+1$ where $dist_T(x_1, z_1) \geqslant 1$. 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. □

Lemma 4.4. *Assume that a class $\mathcal{G}$ of graphs has a simple* CMSO$_1$ *interpretation I in a class $\mathcal{T}_d$ of finite coloured rooted trees of height at most d. Then there* *exists m such that the following holds: every graph $G \in \mathcal{G}$, where $G = I(T)$ for some $T \in \mathcal{T}_d$, has an m-coloured tree-model U of depth d, such that the rooted tree U is obtained from T by "growing leaves" from those nodes of T that belong to the domain of I and have distance less than d from the root.*

Specially, if all vertices in the domain of I are leaves of T at distance d from the root, then $U = T$.

Here the operation of *growing a leaf* from an internal node u of a rooted tree T of height d means to add a new branch (a path) from u to a new leaf u' such that the distance from the root to u' is exactly d. Only one new leaf will be grown from an internal node. See Figure 6.

Proof. Let the simple CMSO_1 interpretation $I(\mathcal{T}_d) = \mathcal{G}$ be given by the formulas $I = (\alpha, \beta)$. Since α, β are finite formulas, they can "see" at most some m' of the colours of $\mathcal{T}_d$. We may choose m' such that the least common multiple of the b values of all $\bmod_{a,b}$ predicates in α, β is also at most m'. Recalling the definition of a simple interpretation, every $G \in \mathcal{G}$ is interpreted in some m'-coloured tree $T_G \in \mathcal{T}_d$ as follows: $V(G) = \{x \in V(T_G) : T_G \models \alpha(x)\}$ and $E(G) = \{xy : x, y \in V(G) \wedge T_G \models \beta(x, y)\}$.

For technical reasons, we transform β into a closed sentence, $\beta' \equiv \exists x, y\big(L(x) \wedge L(y) \wedge \beta(x, y)\big)$, where L is a new label (added to existing colours of nodes of the tree). Later, we will add the label L to precisely two nodes of T_G for which we will need to test adjacency in G.

Let $G \in \mathcal{G}$ be a fixed graph and let $T = T_G$, as above. Let q be the number of quantifiers in β', and for R from Lemma 4.2, let $R'(i) = R(q, m', i) + 2$. Now the tree T is R'-reduced to $T_0 \subseteq T$.

Suppose that u, v is a pair of nodes of T for which we want to test adjacency in G. Let $T[L(u), L(v)]$ denote the tree T in which the label L has been added precisely to some two $u, v \in V(T)$. We correspondingly denote by $T_0[L(u), L(v)]$ the reduced tree as described above. Note that, by automorphism, we may always assume $u, v \in V(T_0)$. Furthermore, while forming $T_0[L(u), L(v)]$, we only remove such components from colour-preserving isomorphism classes of $T_w[L(u), L(v)] - w$ which are of size greater than $R'_i - 2 = R_i$; thereby accounting for the possibility that some (at most two) components of $T_w[L(u), L(v)] - w$ have received the label L. Henceforth, it follows by repeated application of Lemma 4.2 that $T[L(u), L(v)] \models \beta' \iff T_0[L(u), L(v)] \models \beta'$.

Consequently, for each pair u, v, one can determine whether or not it forms an edge in G simply by testing if T_0 with a suitable assignment of L satisfies β'. With this crucial finding at hand, we now easily obtain a tree-model U for G of depth d and a number of colours m depending only on bounded-size T_0.

For $w \in V(T)$, we denote by $h(w)$ the distance from w to the root of T. Starting with the rooted Steiner tree of $V(G)$ in $T = T_G$, we construct U by growing leaves from all the nodes w of T such that $w \in V(G)$ and $h(w) < d$, in order to literally satisfy Definition 3.1. The newly grown leaf w' will now interpret the corresponding vertex of G instead of original w, that is, we will identify $w \in V(G)$ with $w' \in V(U)$, and set $h'(w) := d - h(w)$. Then, for each

$v \in V(G)$, let $Or(v)$ be the automorphism orbit of v in T_0. We will assign the colour $\langle T_0, Or(v), h'(v)\rangle$ to v.

What remains is to prove that, for any $u, v \in V(G) \subseteq V(U)$, only the height of the least common ancestor $z = u \wedge v$ in U and the colours $\langle T_0, Or(u), h'(u)\rangle$ and $\langle T_0, Or(v), h'(v)\rangle$ are sufficient to determine whether $T \models \beta(u, v)$.

For any pair $u, v \in V(G)$, we simply determine $dist_T(u, z)$ and $dist_T(v, z)$ (note that here we mean the distance in T and not in U) from $h'(u), h'(v)$ and the height of $z = u \wedge v$ in U. Using these distances and the orbits $Or(u), Or(v)$ we can hence, by Lemma 4.3, determine the position of the pair u, v within T_0 up to a colour-preserving automoprhism. So, rephrasing, we can construct $T_0[L(u), L(v)]$ only from the height of $z = u \wedge v$ and the colours of u, v in U. Since it is a finite task to decide whether $T_0[L(u), L(v)] \models \beta'$, we finish by previous $T_0[L(u), L(v)] \models \beta' \iff T[L(u), L(v)] \models \beta' \iff T \models \beta(u, v)$. □

4.2 Stability under transductions

The first important consequence of Theorem 4.1 is that the shrub-depth of a graph class is preserved under non-copying CMSO_1 transductions.

Theorem 4.5. *Let $d \geqslant 1$ be an integer, $\mathcal{G}$ be a graph class of shrub-depth d, and τ be a non-copying* CMSO_1 *transduction. Then the shrub-depth of the transduction image $\tau(\mathcal{G})$ is again at most d.*

Proof. Let $\mathcal{G} \subseteq \mathcal{TM}_m(d)$, and let I_1 denote the corresponding interpretation of $\mathcal{G}$ in a class of m-coloured tree-models of depth d. Assume $\tau = \tau_0 \circ \varepsilon$ where τ_0 is a basic transduction and ε is a p-parameter expansion. Since each of the p parameters can be encoded by a binary label added to the above m colours, we have got that $\varepsilon(\mathcal{G})$ has an interpretation I_1' in a class $\mathcal{T}$ of $(2^p m)$-coloured rooted trees of height d. Let I_0 be the simple CMSO_1 interpretation underlying τ_0. Then $\tau(\mathcal{G}) \subseteq I_0\big(I_1'(\mathcal{T})\big)$. Since $I_0 \circ I_1'$ is again a CMSO_1 interpretation, the latter class has shrub-depth at most d by Theorem 4.1 and the claim follows. □

We now look at the more general case of copying CMSO_1 transductions. One cannot immediately extend Theorem 4.5 towards this case since, for example, a 2-copying transduction of the class of edge-less graphs (shrub-depth 1) contains all perfect matchings (shrub-depth 2). This is, however, only a technical problem which we resolve simply by allowing "copying" tree-models here.

Informally, a *k-copied tree-model* is a tree-model T as in Definition 3.1, with an exception that every leaf of T holds an ordered $\leqslant k$-tuple of distinct vertices of G and the existence of an edge can depend also on the tuple of a vertex and its index within the tuple. This is formally stated (with a twist) as follows:

Definition 4.6 (k-copied tree-model). *A graph G has a k-*copied tree-model *of m colours and* "depth" *d if G has an ordinary tree-model T of m colours and depth $d + 1$ such that every node of T at distance d from the root has at most k descendants (the leaves). The class of all graphs G having such a k-copied tree-model is denoted by $\mathcal{TM}_m^k(d)$.*

A class of graphs $\mathcal{G}$ has copying shrub-depth d *if there exist m, k such that $\mathcal{G} \subseteq \mathcal{TM}_m^k(d)$, while for all natural m', k' it is $\mathcal{G} \nsubseteq \mathcal{TM}_{m'}^{k'}(d-1)$.*

Notice that $\mathcal{TM}_m^1(d) = \mathcal{TM}_m(d)$, but this is not true in general for higher values of k. One can easily observe that every graph class of copying shrub-depth d is contained in a suitable k-copying transduction of a class of ordinary shrub-depth d. We complement this observation with:

Theorem 4.7. *Let $d \geqslant 1$, $\mathcal{G}$ be a graph class of copying shrub-depth d, and τ be a* CMSO_1 *transduction. Then the copying shrub-depth of $\tau(\mathcal{G})$ is again at most d.*

Proof. Let $\tau = \tau_0 \circ \gamma \circ \varepsilon$ where τ_0 is a basic CMSO_1 transduction, γ is a k-copy operation and ε is a p-parameter expansion.

We remark that, thanks to transitivity of transductions, it is enough to prove this statement in the case that $\mathcal{G}$ is of ordinary shrub-depth d. So, as in the proof of Theorem 4.5, $\mathcal{G} \subseteq \mathcal{TM}_{m_1}(d)$, and let I_1 denote an MSO_1 interpretation of $\mathcal{G}$ in a suitable class $\mathcal{U}$ of m_1-coloured tree-models of depth d. Then, again as before, we can say that $\varepsilon(\mathcal{G})$ has an interpretation I_1' in the corresponding class $\mathcal{U}'$ of $(2^p m_1)$-coloured tree-models of depth d, that is, $\varepsilon(\mathcal{G}) = I_1'(\mathcal{U}')$.

Next, we present an alternative view of $\gamma(I_1'(\mathcal{U}'))$. In the coming arguments it is important that the domain of I_1' (which is to be copied) is restricted only to leaves of the trees of $\mathcal{U}'$. For $U \in \mathcal{U}'$, let U^+ be the $(k2^p m_1)$-coloured tree-model of depth $d{+}1$ constructed as follows: for each leaf u of U of colour c, add k new descendant leaves with the parent u and of distinct colours $(c, 1), \ldots, (c, k)$. (Actually, U^+ is also a k-copied tree-model of depth d according to Definition 4.6.) Let $\mathcal{U}^+ = \{U^+ : U \in \mathcal{U}'\}$. From the definition of k-copy γ (in Definition 2.6), one can now easily come up with a basic transduction τ_2 such that $\gamma(I_1'(\mathcal{U}')) = \tau_2(\mathcal{U}^+)$.

Recall that $\tau(\mathcal{G}) = (\tau_0 \circ \gamma)(I_1'(\mathcal{U}')) = (\tau_0 \circ \tau_2)(\mathcal{U}^+)$. If I denote the simple CMSO_1 interpretation underlying basic $\tau_0 \circ \tau_2$, then $\tau(\mathcal{G}) \subseteq I(\mathcal{U}^+)$. By Lemma 4.4, there exists m such that every graph $H \in I(\mathcal{U}^+)$, where $H = I(U_1)$ for some $U_1 \in \mathcal{U}^+$, has an m-coloured tree-model U_2 of depth $d + 1$. Since, moreover, the domain of I is restricted to the leaves of U_1 (recall τ_2), we have got $U_2 = U_1$ as rooted trees. This, in particular, means that U_2 is also a k-copied tree-model of m colours and "depth" d of the graph H. Consequently, $\tau(\mathcal{G})$ has copying shrub-depth at most d. □

4.3 On MSO_1-transduction hierarchy

The second interesting consequence of Theorems 4.1 and 4.7 claims that every graph class of bounded shrub-depth "falls under" precisely one of the integer values of copying shrub-depth according to transduction equivalence (both MSO_1 and CMSO_1). This coming result is tightly related to the main result of Blumensath and Courcelle in [1] – a complete characterization of the MSO_2-transduction hierarchy, and it partially answers the similar question about MSO_1-transduction hierarchy also from [1].

We start with some necessary technical terms. Fix a logical language of transductions (such as MSO_1 or CMSO_1 of simple undirected graphs). For two classes of relational structures (graphs in our case) $\mathcal{K}, \mathcal{L}$, we write $\mathcal{K} \sqsubseteq \mathcal{L}$ if there exists a transduction τ such that $\mathcal{K} \subseteq \tau(\mathcal{L})$. Similarly we write $\mathcal{K} \sqsubsetneq \mathcal{L}$ if $\mathcal{K} \sqsubseteq \mathcal{L}$ but $\mathcal{L} \not\sqsubseteq \mathcal{K}$, and $\mathcal{K} \equiv \mathcal{L}$ if both $\mathcal{K} \sqsubseteq \mathcal{L}$ and $\mathcal{L} \sqsubseteq \mathcal{K}$ hold true.

The relation $\sqsubseteq$ forms a quasi-ordering on the considered classes of structures, as can be easily seen [1]. The research question here is to describe the underlying ordering. This has been done in full in [1] for classes of undirected graphs and MSO_2 transductions (precisely, for the vertex-edge incidence structures of undirected graphs). The same question for MSO_1 transductions has been left widely open in [1], and here we provide the partial answer for all graph classes of bounded shrub-depth as follows.

Theorem 4.8. *Let $\mathcal{T}_d$ denote the class of all finite rooted trees of height at most d. In the scope of either* MSO_1 *or* CMSO_1 *transductions, the following holds*

$$\mathcal{T}_1 \sqsubsetneq \mathcal{T}_2 \sqsubsetneq \mathcal{T}_3 \sqsubsetneq \ldots \sqsubsetneq \mathcal{T}_d \ldots .$$

Moreover, for any graph class $\mathcal{G}$ of bounded shrub-depth there is an integer d such that $\mathcal{G} \equiv \mathcal{T}_d$.

Before getting to the proof, we first comment on this claim. There are basically two sides of Theorem 4.8:

- Strict containment; we have $\mathcal{T}_i \sqsubsetneq \mathcal{T}_{i+1}$ for all $i \geqslant 1$. This follows already from [1].
- Completeness of the hierarchy; there are no other classes of bounded shrub-depth than those equivalent to some $\mathcal{T}_d$. This does not seem to follow from [1] in any way, and so here we provide a proof which is based on our results about copying shrub-depth and a translation of some of the arguments of [1].

For the latter we first establish two supplementary lemmas. Let $\overline{T}_d^r$ denote the complete rooted r-ary tree of height d.

Lemma 4.9. *Let $\mathcal{G}$ be a graph class of bounded shrub-depth. If there exist integers d, m, r such that every graph $G \in \mathcal{G}$ has an m-coloured tree-model of depth d not containing $\overline{T}_d^r$ as a rooted subtree, then the copying shrub-depth of $\mathcal{G}$ is at most $d-1$.*

Proof. Let G have a tree-model with the underlying rooted tree U of height d such that $\overline{T}_d^r \not\subseteq U$. We have borrowed the following high-level proof idea from [1, Lemma 4.12].

Let $R \subseteq V(U)$ be the minimal (by inclusion) set of nodes such that R contains all the leaves of U, and R contains every internal node of U which has at least r of its children in R. Let $F \subseteq E(U)$ be the set of edges having one end in R and the other in $V(U)\backslash R$. The root of U is not in R since $\overline{T}_d^r \not\subseteq U$. So, every root-to-leaf path in U contains an edge from F. Moreover, every internal node of U is incident with at most $r-1$ edges of F (or it would be added to R).

Now, to every non-leaf edge $f \in F$ with parent end v we assign a label $\ell_f = (i, j)$, where $0 \leqslant i \leqslant d-2$ is the distance of v from the root and $1 \leqslant j < r$

is the index of f among all F-edges incident with v (in an arbitrary fixed ordering of the children). Then, in the subtree U_f below f in U, we subdivide all the leaf edges of U_f and we add the label ℓ_f to (the colours of) the leaves of U_f. Then we contract f. Let U' denote the resulting labelled tree (which is again of height d). One can routinely verify that information additionally provided by the added labels (ℓ_f) is sufficient for U' to be a tree-model of G, too. Furthermore, our construction of U' guarantees that U' actually is an $(r-1)$-copied tree-model of depth $d-1$, as in Definition 4.6. Since this holds, with the same d, r, for every $G \in \mathcal{G}$, the copying shrub-depth of $\mathcal{G}$ is at most $d-1$. $\square$

Lemma 4.10. *For every integers $d, m \geqslant 1$ there exists a non-copying* MSO_1 *transduction $\sigma_{d,m}$ such that the following holds: if, for an integer r and a graph $G \in \mathcal{TM}_m(d)$, every m-coloured tree-model of depth d of G contains the tree $\overline{T}_d^{\,r}$ as a rooted subtree, then $\overline{T}_d^{\,r-1} \in \sigma_{d,m}(G)$.*

We remark that Blumensath and Courcelle [personal communication] have established a statement similar to Lemma 4.10, but it has not been published. For the sake of completeness, we give our independent proof here.

Proof. Our strategy is to construct a very specific tree-model U of G, such that we can interpret in suitably labelled G a tree $U' \subseteq U$ which is "nearly U" in the sense that only one child of each node of the underlying tree of U is missing (it is used to represent this node instead). From the assumption $\overline{T}_d^{\,r} \subseteq U$ we can then conclude that $\overline{T}_d^{\,r-1}$ will be contained in the respective non-copying transduction image $\sigma_{d,m}(G)$ of this interpretation.

We use technical terms from [16]. Assume T is a tree-model of G, with an internal node u, and let W be the set of leaves of T_u. We say that a tree-model T' is obtained from T by splitting T_u along $X \subseteq W$ if a disjoint copy T'_u of T_u with the same parent is added into T, and then T_u is restricted to the leaves in $W \backslash X$ while T'_u is restricted to those in the corresponding copy of X. A tree-model T is *unsplittable* if no such splitting T' of T represents the same graph G as T does.

Fix now an unsplittable tree-model U of G (which obviously exists, by repeated splits). Let U be Q-reduced to $U_0 \subseteq U$ for $Q \equiv 2$ (cf. Section 4.1 for "reduced"), where U_0 is of constant size depending on d, m. We colour the vertices of G by their colours in T, and additionally give individual distinguishing labels to those (constantly many) vertices which are the leaves of U_0.

- By [16, Lemma 5.10], there exists an FO-definable relation $\sim$ (depending on U_0) on the vertices not in U_0 such that $G \models x \sim y$ if, and only if, x, y are leaves of the same component (subtree) of $U - V(U_0)$.

From $\sim$ one can recursively construct FO-definable relations $\approx_i$ on $V(G)$, for $i = 1, \ldots, d$, such that the following holds: $G \models x \approx_i y$ if, and only if, there is a node w of U such that U_w is of height i and $x, y \in V(U_w)$. The precise technical details are analogous to the proof of [16, Theorem 5.14], and we refrain from repeating them here.

Finally, from each equivalence class of $\approx_1$ we choose an arbitrary representative, and give all these representatives a new label ν_1. Recursively, from each equivalence class of $\approx_i$, $i \geqslant 2$, we choose a representative among those labelled ν_{i-1}, and give them an additional label ν_i. We can now easily interpret the desired tree U' in G using the relations $\approx_i$ and the labels ν_i. Consequently, since $\overline{T}_d^{r-1} \subseteq U'$ this gives $\sigma_{d,m}$ such that $\overline{T}_d^{r-1} \in \sigma_{d,m}(G)$. $\square$

We can now finish the main result.

Proof (of Theorem 4.8). The relation $\mathcal{T}_1 \sqsubseteq \mathcal{T}_2 \sqsubseteq \mathcal{T}_3 \dots$ is trivial. As for the strictness side, $\mathcal{T}_i \sqsubsetneq \mathcal{T}_{i+1}$ for all $i \geqslant 1$ has been proved in [1] also in the scope of MSO_1 transductions (besides MSO_2). The same holds in the scope of CMSO_1, e.g., by Theorem 4.1 since tree-models do not involve any counting.

We are left with proving completeness of our hierarchy. Let us consider any graph class $\mathcal{G}$ of shrub-depth d, and let m be such that $\mathcal{G} \subseteq \mathcal{TM}_m(d)$. Since the shrub-depth of $\mathcal{G}$ is not $d-1$, by Lemma 4.9 we obtain that for every integer r there exists $G_r \in \mathcal{G}$ such that, every m-coloured tree-model of G_r of depth d contains $\overline{T}_d^r$. Then, by Lemma 4.10, there is an MSO_1 transduction $\sigma_{d,m}$ such that $\overline{T}_d^{r-1} \in \sigma_{d,m}(G)$. Hence, $\overline{\mathcal{T}}_d := \{\overline{T}_d^s : s \in \mathbb{N}\} \subseteq \sigma_{d,m}(\mathcal{G})$. Since $\mathcal{T}_d$ is easily a transduction of $\overline{\mathcal{T}}_d$, we conclude that $\mathcal{G} \equiv \mathcal{T}_d$. $\square$

5 Concluding notes

The structural properties of classes of bounded shrub-depth, in Section 3, leave one important question widely open: what is a nice asymptotic structural characterization of graph classes of unbounded shrub-depth? There are indications, related to matroid theory and to the notion of rank-depth by DeVos, Kwon and Oum, that the following might be the ultimate answer:

- [23, Conjecture 6.3] A class $\mathcal{C}$ of graphs is of bounded shrub-depth if, and only if, there exists an integer t such that no graph $G \in \mathcal{C}$ contains a path of length t as a vertex-minor.

On the other hand, in relation to the transduction hierarchy studied in Section 4, the following seems a plausible conjecture:

Conjecture 5.1. A class $\mathcal{C}$ of graphs is of bounded shrub-depth if, and only if, for every CMSO_1 transduction τ there exists an integer t such $P_t \notin \tau(\mathcal{G})$.

While the 'only if' direction follows from Theorems 3.7 and 4.7, the 'if' direction can bee seen as a weaker form of [23, Conjecture 6.3] since a vertex-minor can be captured by a non-copying CMSO_1 transduction.

Finally, we briefly mention a natural extension of the shrub-depth notion to general relational structures (e.g., digraphs). Regarding Definition 3.1 of a tree-model, the extension is straightforward. For any (finite) signature of a relational structure $\mathcal{S}$, the domain of $\mathcal{S}$ is again the set of leaves of T, and we consider (one of) its k-ary relational symbol R. We state that, for an ordered k-tuple $x_1, \dots, x_k$

from the domain, $R(x_1, \ldots, x_k)$ depends only on the colours of $x_1, \ldots, x_k$ and the shape of the rooted Steiner tree of the leaves $x_1, \ldots, x_k$. Hence we can define the shrub-depth as in Definition 3.3 for any class of relational structures of a given finite signature. (Notice, though, that SC-depth does not extend this way.)

With the previous definition, one may readily extend Theorem 4.1 to classes of relational structures of a fixed finite signature. In fact, it is enough to provide a corresponding extension of technical Lemma 4.3, and the rest of the arguments go smoothly through. The question of the lower levels of the MSO-transduction hierarchy, as in Theorem 4.8, of such classes is left for future investigation.

References

1. A. Blumensath and B. Courcelle. On the monadic second-order transduction hierarchy. *Log. Methods Comput. Sci.*, 6(2), 2010.
2. A. Blumensath and B. Courcelle. Monadic second-order definable graph orderings. *Logical Methods in Computer Science*, 10(1), 2014.
3. H. L. Bodlaender, J. S. Deogun, K. Jansen, T. Kloks, D. Kratsch, H. Müller, and Z. Tuza. Rankings of graphs. *SIAM J. Discrete Math.*, 11(1):168–181, 1998.
4. B. Courcelle. The monadic second order logic of graphs I: Recognizable sets of finite graphs. *Inform. and Comput.*, 85:12–75, 1990.
5. B. Courcelle and J. Engelfriet. *Graph Structure and Monadic Second-Order Logic: a Language-Theoretic Approach*, volume 138 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, 2012.
6. B. Courcelle, J. Engelfriet, and G. Rozenberg. *Graph Grammars and Their Application to Computer Science: 4th International Workshop Bremen, Germany, March 5–9, 1990 Proceedings*, chapter Context-free handle-rewriting hypergraph grammars, pages 253–268. Springer Berlin Heidelberg, Berlin, Heidelberg, 1991.
7. 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.
8. B. Courcelle and S. Olariu. Upper bounds to the clique width of graphs. *Discrete Appl. Math.*, 101(1-3):77–114, 2000.
9. J. Deogun, T. Kloks, D. Kratsch, and H. Müller. On vertex ranking for permutation and other graphs. In *STACS'94*, volume 775 of *LNCS*, pages 747–758. Springer, 1994.
10. R. Diestel. *Graph Theory*, volume 173 of *Graduate texts in mathematics*. Springer, New York, 2005.
11. G. Ding. Subgraphs and well-quasi-ordering. *J. Graph Theory*, 16(5):489–502, 1992.
12. M. Elberfeld, M. Grohe, and T. Tantau. Where first-order and monadic second-order logic coincide. In *LICS'12*, pages 265–274, 2012.
13. 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.
14. 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.
15. J. Gajarský. Efficient solvability of graph MSO properties. Master's thesis, Masaryk University, Brno, 2012.

16. J. Gajarský and P. Hliněný. Kernelizing MSO properties of trees of fixed height, and some consequences. *Logical Methods in Computer Science*, 11(1), 2015.
17. R. Ganian. Thread graphs, linear rank-width and their algorithmic applications. In *IWOCA'10*, volume 6460 of *LNCS*, pages 38–42. Springer, 2010.
18. R. Ganian. Improving vertex cover as a graph parameter. *Discrete Mathematics & Theoretical Computer Science*, 17(2):77–100, 2015.
19. R. Ganian, P. Hliněný, J. Nešetřil, J. Obdržálek, P. O. de Mendez, and R. Ramadurai. When trees grow low: Shrubs and fast MSO_1. In *MFCS'12*, volume 7464 of *LNCS*, pages 419–430. Springer, 2012.
20. 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.
21. 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.
22. P. Heggernes, D. Meister, and C. Papadopoulos. Characterising the linear clique-width of a class of graphs by forbidden induced subgraphs. *Discrete Appl. Math.*, 160(6):888–901, 2012.
23. P. Hliněný, O. Kwon, J. Obdržálek, and S. Ordyniak. Tree-depth and vertex-minors. *Eur. J. Comb.*, 56:46–56, 2016.
24. P. Hliněný and S. Oum. Finding branch-decomposition and rank-decomposition. *SIAM J. Comput.*, 38:1012–1032, 2008.
25. Ö. Johansson. Clique-decomposition, NLC-decomposition, and modular decomposition — relationships and results for random graphs. *Congr. Numer.*, 132:39–60, 1998.
26. M. Lampis. Algorithmic meta-theorems for restrictions of treewidth. *Algorithmica*, 64(1):19–37, 2012.
27. M. Lampis. Model checking lower bounds for simple graphs. *Logical Methods in Computer Science*, 10(1), 2014.
28. J. Nešetřil and P. Ossona de Mendez. Grad and classes with bounded expansion I. Decompositions. *European J. Combin.*, 29(3):760–776, 2008.
29. J. Nešetřil and P. Ossona de Mendez. *Sparsity (Graphs, Structures, and Algorithms)*, volume 28 of *Algorithms and Combinatorics*. Springer, 2012. 465 pages.
30. J. Nešetřil and P. Ossona de Mendez. Tree-depth, subgraph coloring and homomorphism bounds. *European J. Combin.*, 27(6):1024–1041, 2006.
31. S. Oum and P. D. Seymour. Approximating clique-width and branch-width. *J. Combin. Theory Ser. B*, 96(4):514–528, 2006.
32. N. Robertson and P. D. Seymour. Graph minors. II. Algorithmic aspects of tree-width. *J. Algorithms*, 7(3):309–322, 1986.
33. P. Schaffer. Optimal node ranking of trees in linear time. *Inform. Process. Lett.*, 33:91–96, 1989/90.
34. E. Wanke. k-NLC graphs and polynomial algorithms. *Discrete Appl. Math.*, 54:251–266, 1994.