Kernelizing MSO Properties of Trees of Fixed Height, and Some Consequences *

Jakub Gajarský and Petr Hliněný **

Faculty of Informatics, Masaryk University
Botanická 68a, Brno, Czech Republic
\{xgajar,hlineny\}@fi.muni.cz

Abstract. Fix an integer \( h \geq 1 \). In the universe of coloured trees of height at most \( h \), we prove that for any MSO formula with \( r \) variables there exists a set of kernels, each of size bounded by an elementary function of \( r \) and the number of colours. This yields two noteworthy consequences. Consider any graph class \( \mathcal{G} \) having a simple MSO interpretation in the universe of coloured trees of height \( h \) (equivalently, \( \mathcal{G} \) is a class of shrub-depth \( h \)). First, \( \mathcal{G} \) admits an MSO model checking algorithm whose runtime has an elementary dependence on the formula size. Second, on \( \mathcal{G} \) the expressive powers of FO and MSO coincide (which extends a 2012 result of Elberfeld, Grohe, and Tantau [9]).

Keywords: FO logic; MSO logic; graph; tree-depth; shrub-depth

1 Introduction

First order (FO) and monadic second-order (MSO) logics play undoubtedly crucial role in computer science. Besides traditional tight relations to finite automata and regular languages, this is also witnessed by their frequent occurrence in the so called algorithmic metatheorems which have gained increasing popularity in the past few years. The term algorithmic metatheorem commonly refers to a general algorithmic toolbox ready to be applied onto a wide range of problems in specific situations, and MSO or FO logic is often used in the expression of this “range of problems”.

One of the perhaps most celebrated algorithmic metatheorems (and the original motivation for our research) is Courcelle’s theorem [2] stating that every graph property \( \phi \) expressible in the MSO\(_2\) logic of graphs (allowing for both vertex and edge set quantifiers) can be decided in linear FPT time on graphs of bounded tree-width. Courcelle, Makowsky, and Rotics [4] then have analogously addressed a wider class of graphs, namely those of bounded clique-width, at the expense of restricting \( \phi \) to MSO\(_1\) logic (i.e., with only vertex set quantification).

Regarding Courcelle’s theorem [2] and closely related [11], it is worth to remark that a solution can be obtained via interpretation of the respective

* Preliminary extended abstract in FSTTCS 2012.
** Supported by the Czech Science Foundation; project P202/11/0196.
graph problem into an MSO formula over coloured trees (which relates the
topic all the way back to Rabin’s S2S theorem [19]). However, a common
drawback of these metatheorems is that, when their runtime is expressed as
\[ O\left(f(\phi, width(G)) \cdot |G|\right), \]
where the height of width(G) \(a\) depends on \(\phi\), precisely on the quantifier alternation depth of \(\phi\) (i.e., \(f\) is a non-elementary function of the parameter \(\phi\)). The latter is not surprising since Frick and Grohe [10] proved that it is not possible to avoid a non-elementary tower of exponents in deciding MSO properties on all trees or coloured paths (unless P=NP), and Lampis [16] proved an analogous negative result even for uncoloured paths (unless EXP=NEXP).

The aforementioned negative results leave main room for possible improve-
ment on suitably restricted subclass(es) of all coloured trees, namely on those
avoiding long paths. In this respect, our first result (Theorem 3.2) gives a new
algorithm for deciding MSO properties \(\phi\) of rooted coloured trees \(T\) of fixed
height \(h\). The algorithm (Corollary 3.4) uses so called kernelization—which
means it efficiently reduces the input tree into an equivalent one (the kernel)
of elementarily bounded size; by
\[ 2^{O(|\phi|^2)} h. \]

In the complexity aspects, our result “trades” quantifier alternation depth of \(\phi\)
from Courcelle’s theorem for bounded height of the tree. We refer to Section 3
for details and exact expression of runtime as well as for an extension to counting
MSO logic.

From a more general perspective our algorithm can be straightforwardly ap-
plied to any suitable “depth-structured” graph class via efficient interpretability
of logic theories, such as to graph classes of bounded tree-depth or of bounded
shrub-depth [12]. This (asymptotically) includes previous results of Lampis [16]
and Ganian [11] as special cases. Even more, the scope of our result can be ex-
tended to the so called LinEMSO optimization and enumeration framework, see
e.g. in [4], over such graph classes as follows: the algorithmic metatheorems of
[2114] can be treated using finite tree-automata (with non-elementary numbers
of states in general), however, in the universe of a suitable “depth-structured”
graph class only very few of the automaton states correspond to some of the
kernels and so are actually reachable. This in Section 4 concludes the first half
of our paper.

There are also other sides of the main result. First, the initial discovery of
Theorem 3.2 was the prime motivation for defining shrub-depth in [12], and
a key ingredient in the proof that shrub-depth is stable under simple MSO_{1}
interpretations, again in [12].

Second, Elberfeld, Grohe, and Tantau [9] proved in 2012 that FO and MSO_{2}
have equal expressive power on the graphs of bounded tree-depth. Having The-
orem 3.2 at hand, we can provide a relatively simple alternative proof of this
result. Furthermore, using some more sophisticated combinatorial tools, namely
well-quasi-ordering, we prove a new result (Theorem 5.13) that FO and MSO
have equal expressive power on any graph class of bounded shrub-depth. Unfor-
tunately, due to lack of a suitable “forbidden substructure” characterization of
shrub-depth, we are not yet able to prove the converse direction, but we conjec-
ture that a hereditary class on which FO and MSO coincide must have bounded
shrub-depth (Conjecture 6.1).

2 Definitions

We assume standard terminology and notation of graph theory, see e.g. Dies-
tel [5]. Our graphs are finite, simple and undirected by default. When dealing
with trees, we implicitly consider them as rooted (with the implicit parent-child
tree order) and with unordered descendants. We use \( \subseteq \) for the usual subgraph
relation and \( \subseteq \) for induced subgraphs. A graph class \( \mathcal{C} \) is hereditary
if it is closed under induced subgraph relation, i.e., \( G \in \mathcal{C} \) and \( H \subseteq_i G \) implies
\( H \in \mathcal{C} \).

A class \( \mathcal{C} \) of graphs is well-quasi-ordered (WQO) under \( \preceq \) if, for any infinite
sequence \( (G_1, G_2, \ldots) \subseteq \mathcal{C} \), it is \( G_i \preceq G_j \) for some \( i < j \). If \( \Pi \) is a property
preserved under well-quasi-ordered \( \preceq \) (i.e., if \( G \in \mathcal{C} \) has the property \( \Pi \), then
every \( G' \in \mathcal{C} \) such that \( G' \preceq G \) has \( \Pi \)), then \( \Pi \) is characterized by finitely many
obstacles.

Given a graph \( G \), a tree-decomposition of \( G \) is an ordered pair \( (T, W) \), where \( T \)
is a tree and \( W = \{ W_x \subseteq V(G) \mid x \in V(T) \} \) is a collection of bags (vertex sets
of \( G \)), such that the following hold:

1. \( \bigcup_{x \in V(T)} W_x = V(G) \); 
2. for every edge \( e = uv \) in \( G \), there exists \( x \in V(T) \) such that \( u, v \in W_x \);
3. for each \( u \in V(G) \), the set \( \{ x \in V(T) \mid u \in W_x \} \) induces a subtree of \( T \).

The width of a tree-decomposition \( (T, W) \) is \( \max_{x \in V(T)} |W_x| - 1 \). The tree-width
of \( G \), denoted \( tw(G) \), is the smallest width of a tree-decomposition of \( G \).

Besides tree-width, another useful width measure of graphs is clique-width of a graph \( G \). This is defined for a graph \( G \) as the smallest number of labels \( k = cw(G) \) such that some labelling of \( G \) can be defined by an algebraic \( k\)-expression using the following operations:

1. create a new vertex with label \( i \);
2. take the disjoint union of two labeled graphs;
3. add all edges between vertices of label \( i \) and label \( j \); and
4. relabel all vertices with label \( i \) to have label \( j \).

Monadic second-order logic (MSO) is an extension of first-order logic (FO)
by quantification over sets. The quantifier rank \( qr(\phi) \) is the nesting depth of
quantifiers in \( \phi \). The formulas of quantifier rank \( 0 \) are called quantifier free.
Counting monadic second-order logic (CMSO) is an extension of MSO which
allows use of predicates \( mod_{a,b}(X) \), where \( X \) is a set variable. The semantics of
the predicate \( mod_{a,b}(X) \) is that the set \( X \) has \( a \) modulo \( b \) elements.
Let $\sigma$ and $\tau$ be relational vocabularies and let $L \in \{\text{FO}, \text{MSO}, \text{CMSO}\}$. A simple interpretation of $\tau$ into $\sigma$ is a tuple $I = (\nu(x), \{\eta_R(\bar{x})\}_{R \in \tau})$ of $L[\sigma]$-formulas where $\nu$ has one free variable and the number of free variables in each $\eta_R$ is equal to the arity of $R$ in $\tau$.

- To every $\sigma$-structure $A$ the interpretation $I$ assigns a $\tau$-structure $A^I$ with the domain $A^I = \{a \mid A \models \nu(a)\}$ and the relations $R^I = \{\bar{a} \mid A \models \eta_R(\bar{a})\}$ for each $R \in \tau$. We say that a class $\mathcal{C}$ of $\tau$-structures has an interpretation into a class $\mathcal{D}$ of $\sigma$-structures if there exists an interpretation $I$ such that for each $C \in \mathcal{C}$ there exists $D \in \mathcal{D}$ such that $C \equiv D^I$, and for every $D \in \mathcal{D}$ the structure $D^I$ is isomorphic to a member of $\mathcal{C}$.

- The interpretation $I$ of $\tau$ into $\sigma$ defines a translation of every $L[\sigma]$-formula $\phi$ to an $L[\tau]$-formula $\phi^I$ as follows:
  - every $\exists x. \phi$ is replaced by $\exists x. (\nu(x) \land \phi^I)$,
  - every $\exists X. \phi$ is replaced by $\exists X. (\forall y (y \in X \rightarrow \nu(y)) \land \phi^*)$, and
  - every occurrence of a $\sigma$-atom $R(\bar{x})$ is replaced by the corresponding formula $\eta_R(\bar{x})$.

We make use of the following claim:

**Lemma 2.1 ([13]).** Let $I$ be interpretation of $\tau$ into $\sigma$. Then for all $L[\tau]$-formulas $\phi$ and all $\sigma$-structures $A$

$$A \models \phi^I \iff A^I \models \phi.$$ 

A great part of our paper deals with MSO logic of rooted trees, which are structures with a single parent-child binary relation. For general graphs, however, there are two established but inequivalent views of them as relational structures. In the one-sorted adjacency model of graphs MSO specifically reads as follows:

**Definition 2.2 (MSO\textsubscript{1} logic of graphs).** The language of MSO\textsubscript{1} contains the expressions built from the following elements:

- variables $x, y, \ldots$ for vertices, and $X, Y, \ldots$ for sets of vertices,
- the predicates $x \in X$ and edge$(x, y)$ with the standard meaning,
- equality for variables, the connectives $\land, \lor, \neg, \rightarrow$, and the quantifiers $\forall, \exists$ over vertex and vertex-set variables.

MSO\textsubscript{2} logic of graphs extends MSO\textsubscript{1} by allowing quantification over edge sets. Formally, one can consider graphs as two-sorted structures (the two sorts being vertex-set and edge-set of a graph) with adjacency and incidence predicates. Although our results are concerned also with MSO\textsubscript{2}, we refrain from giving full definition of MSO\textsubscript{2} logic of graphs, since by the following theorem we can avoid using MSO\textsubscript{2} explicitly.

\textsuperscript{1} Here we slightly deviate from the more common formulation “interpretation from $\sigma$ to $\tau$”. Our definition stresses that we use interpretations algorithmically – we translate input graph into a labelled tree and solve modification of the original formula on it.
Theorem 2.3 (Courcelle [3]). Let $C$ be the class of planar graphs, or of graphs of degree at most $k$ (for any fixed $k$) or of finite graphs of tree-width at most $k$ (for any fixed $k$). A property of graphs in $C$ is MSO$_2$-expressible if, and only if, it is MSO$_1$-expressible.

We often use labelled graphs, i.e. graphs where each vertex can have labels from some finite set $\text{Lab}$. Labels are modelled by unary predicates. Sometimes we refer to colours instead of labels; the colour of a vertex is the combination of its labels (each vertex has exactly one colour, the number of colours is $\sim 2^{\left|\text{Lab}\right|}$).

For an introduction to parameterized complexity we suggest [7]. Here we just recall that a problem $P$ parameterized by $k$, i.e., with an input $\langle x, k \rangle \in \Sigma^* \times \mathbb{N}$, is fixed parameter tractable, or FPT, if it admits an algorithm in time $O(f(k) \cdot |x|^{O(1)})$ where $f$ is an arbitrary computable function.

3 Trees of Bounded Height and MSO

The primary purpose of this section is to prove coming Theorem 3.2 that for any $m$-coloured tree $T$ of constant height $h$ there exists an efficiently computable subtree $T_0 \subseteq T$ (a kernel) such that, for any MSO sentence $\phi$ of fixed quantifier rank $r$, it is $T \models \phi \iff T_0 \models \phi$, and the size of $T_0$ is bounded by an elementary function of $r$ and $m$ (the dependence on $h$ being non-elementary, though).

Particularly, since checking of an MSO property $\phi$ can be easily solved in time $O^*(2^{c|\phi|})$ on a graph with $c$ vertices (in this case $T_0$) by recursive exhaustive expansion of all quantifiers of $\phi$, this gives a kernelization-based elementary FPT algorithm for MSO model checking of rooted $m$-coloured trees of constant height $h$ (Corollary 3.4).

We need a bit more formal notation. The height $h$ of a rooted tree $T$ is the farthest distance from its root, and a node is at the level $\ell$ if its distance from the root is $h - \ell$. For a node $v$ of a rooted tree $T$, we call a limb of $v$ a subtree of $T$ rooted at some child node of $v$. Our rooted trees are unordered, and they “grow top-down”, i.e. we depict the root on the top. We switch from considering $m$-coloured trees to more convenient $t$-labelled ones, the difference being that one vertex may have several labels at once (and so $m \sim 2^{\text{Lab}}$). We say that two such rooted labelled trees are $l$-isomorphic if there is an isomorphism between them preserving the root and all the labels.

For obtaining the desired kernel, we shall use a concept of reducing a (rooted $t$-labelled) tree $T$ as follows.

Definition 3.1 ($f$-reduction, $f$-reduced). Let $f : \mathbb{N} \to \mathbb{N}$ be a function, called a threshold function. Assuming a node $v \in V(T)$ at level $i + 1 > 0$ and a limb $B$ of $v$ such that there exist at least $f(i)$ other limbs of $v$ in $T$ which are

\footnote{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).}
all $l$-isomorphic to $B$; we say that $T$ $f$-reduces in one step to $T - V(B)$. A tree $T$ $f$-reduces to $T_0 \subseteq T$ if there is a sequence of one-step $f$-reductions from $T$ to $T_0$. The tree $T_0$ is $f$-reduced if no further $f$-reduction step is possible.

A straightforward induction shows that any $f$-reduced $t$-labelled rooted tree of fixed height has size bounded with respect to $f$ and $t$.

### 3.1 Reduction to kernel

Considering MSO sentences with $q$ element quantifiers and $s$ set quantifiers, we use the following threshold function $R$ to define our reduced kernel (where $k$ is an arbitrary integer parameter that will count all the labels used in our proof):

\[
R_{q,s,k}(i) = q \cdot N_{q,s,k}(i)^s, \quad \text{where}
\]

\[
N_{q,s,k}(0) = 2^k + 1 \geq 2 \quad \text{and}
\]

\[
N_{q,s,k}(i + 1) = 2^k \cdot (R_{q,s,k}(i) + 1)^{N_{q,s,k}(i)} \leq 2^k \cdot (2q \cdot N_{q,s,k}(i)^s)^{N_{q,s,k}(i)} \quad (2)
\]

Note also that all these values are non-decreasing in the parameters $q, s, k$.

The idea behind our approach can be simplified as follows. Fix $i \geq 0$ and any $a \geq R_{q,s,k}(i)$ where $k = t + 3q + s$, and choose an arbitrary $R_{q,s,k}$-reduced rooted $t$-labelled tree $U$ of height $i$. Then no MSO sentence with $q$ element variables and $s$ set variables can distinguish between $a$ disjoint copies and $a + 1$ disjoint copies of $U$. Our full result then reads:

**Theorem 3.2.** Let $T$ be a rooted $t$-labelled tree of height $h$, and let $\phi$ be an MSO sentence with $q$ element quantifiers and $s$ set quantifiers. Suppose that $u \in V(T)$ is a node at level $i + 1$ where $i < h$.

a) If, among all the limbs of $u$ in $T$, there are more than $R_{q,s,t+3q+s}(i)$ pairwise $l$-isomorphic ones, then let $T' \subseteq T$ be obtained by deleting one of the latter limbs from $T$. Then, $T \models \phi \iff T' \models \phi$.

b) Consequently, the tree $T$ $R_{q,s,t+3q+s}$-reduces to $T_0 \subseteq T$ such that $T_0$ itself is $R_{q,s,t+3q+s}$-reduced, and $T \models \phi \iff T_0 \models \phi$.

c) The tree $T_0$ can be computed in linear time (non-parameterized) from $T$, and its size is bounded by

\[
|V(T_0)| \leq \exp^{(h)} \left[ (2^{h+5} - 12) \cdot (t + q + s)(q + s) \right].
\]

Here, $\exp^{(i)}(x)$ denotes the $i$-fold exponential function defined inductively as follows: $\exp^{(0)}(x) = x$ and $\exp^{(i+1)}(x) = 2^{\exp^{(i)}(x)}$. Note that $\exp^{(h)}(x)$ is an elementary function of $x$ for each particular height $h$.

In the case of FO logic, a statement analogous to Theorem 3.2 can be obtained using folklore arguments of finite model theory (even full recursive expansion of all $q$ vertex quantifiers in $\phi$ could “hit” only bounded number of limbs of $u$ and the rest would not matter). However, in the case of MSO logic there are additional nontrivial complications which require new ideas (in addition to
standard tools) in the proof. Briefly saying, one has to recursively consider the internal structure of the limbs of $u$, and show that even an expansion of a vertex-set quantifier in $\phi$ does not effectively distinguish too many of them (and hence some of them remain irrelevant for the decision whether $T \models \phi$).

Before proceeding with formal proof of Theorem 3.2, we need to clarify the meaning of the values $N$:

**Lemma 3.3.** For any natural $i, q, s, k$, there are at most $N_{q,s,k}(i)$ pairwise non-l-isomorphic $R_{q,s,k}$-reduced rooted $k$-labelled trees of height $\leq i$.

**Proof.** This claim readily follows from (1) and (2) by induction on $i$. The base case $i = 0$ is trivial, and the count includes also the empty tree. A rooted $k$-labelled tree $T$ of height $\leq i + 1$ can be described by the label of its root $r$ ($2^k$ possibilities), and the set of its limbs, each one of height $\leq i$. This set of limbs can be fully described by the numbers of limbs (between 0 and $R_{q,s,k}(i)$) in every of $\leq N_{q,s,k}(i)$ possible l-isomorphism classes. Hence by (2) we have got at most $N_{q,s,k}(i + 1)$ possible distinct descriptions of $T$. \[\square\]

**Proof (of Theorem 3.2).** For clarity, we start with a proof sketch.

(I) We are going to use a so called “quantifier elimination” approach. That means, assuming $T \models \phi$ $\iff$ $T' \models \phi$, we look at the “distinguishing choice” of the first quantifier in $\phi$, and encode it in the labeling of $T$ (e.g., when $\phi \equiv \exists x.\psi$, we give new exclusive labels to the value of $x$ and to its parent/children in $T$ and $T'$). By an inductive assumption, we then argue that the shorter formula $\psi$ cannot distinguish between these newly labeled $T$ and $T'$, which is a contradiction.

(II) The traditional quantifier elimination approach—namely of set quantifiers in $\phi$, however, might not be directly applicable to even very many pairwise l-isomorphic limbs in $T$ if their size is unbounded. Roughly explaining, the problem is that a single valuation of a set variable on these repeated limbs may potentially pairwise distinguish all of them. Hence additional combinatorial arguments are necessary to bound the size of the limbs in consideration.

(III) Having resolved technical (II), the rest of the proof is a careful composition of inductive arguments using the formula (1) $R_{q,s,k}(i) = q \cdot N_{q,s,k}(i)^s$.

(a) The whole proof goes through by *means of contradiction*. That is, we assume $T \models \phi$ while $T' \models \neg \phi$ (a counterexample to Theorem 3.2a, where $T'$ implicitly depends on the choice of $u$), up to natural symmetry between $\phi$ and $\neg \phi$ in this context. Let $t' = t + 3q + s$. Let $B_1, \ldots, B_p \subseteq T$ where $p > R_{q,s,t'}(i) \geq 1$ be the pairwise l-isomorphic limbs of $u$ in $T$, as anticipated in Theorem 3.2a.

Note that the height of $B_1$ is at most $i$ (but it may possibly be lower than $i$).

So, say, $T' = T - V(B_1)$. We will apply nested induction, primarily targeting the structure of the sentence $\phi$, or simply the value $q + s$. For that we assume

\[\text{This approach has been inspired by [8], though here it is applied in a wider setting of MSO logic.}\]
\[\phi\] in the prenex form, i.e., with a leading section of all quantifiers. If \(q = s = 0\), then \(\phi\) is a propositional formula which evaluates to true or false without respect to \(T\) or \(T'\). Hence we further assume \(q + s > 0\). Note also the little trick with choice of \(t' = t + 3q + s\) which “makes room” for \(T\) in the course of the proof.

(\textit{Minimality setup}) To overcome the complication in \(T\), we have to deal with limbs \(B_1, \ldots, B_p\) of bounded size. So, among all the assumed counterexamples to Theorem \(3.2a\) for this particular \(\phi\) or symmetric \(\neg \phi\), choose one (meaning precisely the choice of \(T\) and \(u\) within it) which minimizes the size of \(B_1\) (same as the sizes of \(B_2, \ldots, B_p\)). This minimality choice actually represents a secondary induction in our proof.

We would like to show that the l-isomorphic limbs \(B_1, \ldots, B_p\) are \(R_{q,s,t'}\)-reduced. Suppose not, and let \(w_k \in V(B_k)\) be a node at level \(j + 1 \leq i\) such that among all the limbs of \(w_k\) in \(B_k\) there are more than \(R_{q,s,t'}(j)\) pairwise l-isomorphic ones, hereafter denoted by \(D_{k,1}, \ldots, D_{k,r}\) where \(r > R_{q,s,t'}(j)\). This choice is made for all \(k = 1, \ldots, p\) symmetrically, i.e., all the subtrees \(B_k = B_k - V(D_{k,1})\) where \(k = 1, \ldots, p\) are pairwise l-isomorphic, too.

We define a sequence of trees by \(U_0 = T\) and \(U_k = U_{k-1} - V(D_{k,1})\) for \(k = 1, \ldots, p\). Recall that \(U_0 \models \phi\). If it ever happened that \(U_{k-1} \models \phi\) but \(U_k \models \neg \phi\), then we would consider \(U_{k-1}\) and \(w_k\) in place of \(T\) and \(u\) above, and hence contradict the choice minimizing \(B_1\) (which would be replaced with smaller \(D_{k,1}\)). We may thus say that \(U_p \models \phi\). We similarly define \(U'_1 = T'\) and \(U'_k = U'_{k-1} - V(D_{k,1})\) for \(k = 2, \ldots, p\) (recall that \(B_1\) has been removed from \(T'\)). With an analogous argument we conclude that \(U'_p \models \neg \phi\).

Note that, now, \(B_1^-, \ldots, B_p^-\) are pairwise l-isomorphic limbs of \(u\) in \(U_p\), and they are strictly smaller than \(B_1\). Since \(U_p' = U_p - V(B_1^-)\), we may have chosen \(U_p\) and \(u\) in place of \(T, u\), again contradicting minimality of \(B_1\) in the choice above. Indeed, the (original) limbs \(B_1, \ldots, B_p\) are \(R_{q,s,t'}\)-reduced in \(T\).

(\textit{Quantifier elimination:} \(\exists x\)) As the main induction step we now “eliminate” the leading quantifier of \(\phi\) as follows. Suppose first that \(\phi \equiv \exists x. \psi\). Let \(a \in V(T)\) be such that \(T[x = a] \models \psi(x)\). Clearly, it can be chosen \(a \notin V(B_1)\) since \(B_1\) is l-isomorphic to other \(B_2, \ldots, B_p\). On the other hand, \(T'[x = b] \not\models \psi(x)\) for all \(b \in V(T')\).

We define a \((t + 3)\)-labelled tree \(T^a\) which results from \(T\) by adding a new label \(L_x\) exclusively to the node \(a\), a new label \(L_{px}\) exclusively to the parent node of \(a\), and \(L_{xx}\) to the child nodes of \(a\). A tree \(T^a' = T^a - V(B_1)\) is formed analogously from \(T'\). Then we translate the formula \(\psi(x)\) with free \(x\) into a closed one \(\psi^x\) as defined next: All label predicates \(L(x)\) in \(\psi(x)\) are simply evaluated as \(L(a)\) over \(T\) (which is the same as over \(T'\)). Any predicate \(x = y\) is replaced with \(L_x(y)\). Finally, all predicates for edges \((x, y)\) and \((y, x)\) in this parent-child order are replaced with \(L_{cx}(y)\) and \(L_{px}(y)\), respectively. It is trivial that \(T[x = a] \models \psi(x) \iff T^a \models \psi^x\), and \(T'[x = a] \not\models \psi(x) \iff T^a' \not\models \psi^x\).

All the limbs \(B_1, \ldots, B_p\) remain pairwise l-isomorphic in \(T^a\) unless, say, \(a \in V(B_p)\). Even in the latter case we anyway obtain, using \(T\), at least \(p - 1 >
A key observation is that \("casting\) the new label at most \(B\) because, for each \(q\), may create at most \(N\).

On the other hand, \(\phi\).

Suppose now that \((Quantifier elimination: \exists X)\) We are finally getting to the heart of the proof. Suppose now that \(\phi \equiv \exists X, \psi\). Let \(A \subseteq V(T)\) be such that \(T[X = A] \models \psi(X)\). On the other hand, \(T'[X = A'] \not\models \psi(X)\) for all \(A' \subseteq V(T')\). We define a \((t + 1)\)-labelled tree \(T^A\) which results from \(T\) by adding a new label \(L_X\), precisely to all members of \(A\). Then we translate the formula \(\psi(X)\) free \(X\) into a closed one \(\psi^X\) by replacing every occurrence of \(y \in X\) with \(L_X(y)\). Trivially, \(T[X = A] \models \psi^X\).

Note again that \(s - 1\) is the number of set quantifiers in \(\psi\), and that the combined parameter \(t + 1 + 3q + (s - 1) = t + 3q + s = t'\) remains the same.

A key observation is that \("casting\) the new label \(L_X\), onto the limbs \(B_1, \ldots, B_p\), may create at most \(N_{q,s,t'}(i)\) \(l\)-isomorphism classes among them. This is simply because, for each \(k = 1, \ldots, p\), the corresponding \(B_k^A\) carries \(t + 1 \leq t'\), labels, it is of the same height as \(B_k\) and \(R_{q,s,t'}\)-reduced, too. Hence, altogether, there are at most \(N_{q,s,t'}(i)\) pairwise non-\(l\)-isomorphic choices for such \(B_k^A\) by Lemma 3.3.

So, among all \(B_1, \ldots, B_p\), there are at least \(p/N_{q,s,t'}(i)\) pairwise \(l\)-isomorphic limbs, and using \([1]\)

\[
\frac{p}{N_{q,s,t'}(i)} > \frac{R_{q,s,t'}(i)}{N_{q,s,t'}(i)} = q \cdot N_{q,s,t'}(i)^{s-1} \geq R_{q,s-1,t'}(i).
\]

For simplicity, let the latter limbs be \(B_1, \ldots, B_{p'}\) where \(p \geq p' > R_{q,s-1,t'}(i)\).

Now we apply the inductive assumption to \(T^A\), \(u\), and \(\psi^X\). Up to symmetry between the limbs, we get \((T^A)' = T^A - V(B_1)\) such that \(T^A \models \psi^X \iff (T^A)' \models \psi^X\). Now we can define \(A' \subseteq V(T')\) as the set of those nodes having label \(L_X\) in \((T^A)'\), and hence \((T^A)' \models \psi^X \iff T'[X = A'] \models \psi(X)\), a contradiction to the initial assumption.

\((Quantifier elimination: \forall)\) Finally, the cases of universal quantifiers in \(\phi\) are solved analogously \((\not\exists \text{ in place of } \forall)\).

\(b)\) This part readily follows by a recursive bottom-up application of \(a)\) to the whole tree \(T\).

\(c)\) \(T_0\) is easily constructed by a natural adaptation of the classical linear-time tree-isomorphism algorithm. Since \(T_0\) is \(R_{q,s,t'}\)-reduced by \(b)\), where \(t' = t + 3q + s\), we can consider \(T_0 \subseteq U_{h,t,q,s,t}\) where \(U_{h,t,q,s,t}\) is the \("maximal\) \(R_{q,s,t'}\)-reduced rooted \(t\)-labelled tree of height \(h\); \(U_{h,t,q,s,t}\) contains (at each level \(j + 1\)) precisely \(R_{q,s,t'}(j)\) limbs of every \(l\)-isomorphism class of rooted \(t'\)-labelled trees of height \(\leq j\).
Therefore, by Lemma 3.3, the number of descendants at each level $j$ of $U_{h,q,s,t}$ is at most $R_{q,s,t}(j-1) \cdot N_{q,s,t'}(j-1)$. The total number of vertices in $U_{h,q,s,t}$ is at most

$$1 + R_{q,s,t}(h-1) \cdot N_{q,s,t'}(h-1) \cdot (1 + R_{q,s,t}(h-2) \cdot N_{q,s,t'}(h-2) \cdot (1 + \ldots)) \leq \prod_{i=0}^{h-1} \left(1 + R_{q,s,t'}(i) \cdot N_{q,s,t'}(i)\right).$$  \hspace{1cm} (3)

The task is now to estimate, by induction on $i$, the value $1 + R_{q,s,t'}(i) \cdot N_{q,s,t'}(i)$ from above by $\exp^{(i+1)} [(6 \cdot 2^i - 2) t'(q + s)]$. Note that $t' \geq q + s \geq 1$.

$$1 + N_{q,s,t'}(0) \cdot R_{q,s,t'}(0) = 1 + q \cdot \left(2^{t'} + 1\right)^{s+1} \leq 2^{q} \cdot 2^{(t'+1)(s+1)} \leq 2^{2^{t'+(s+1)+q}} \leq 2^{4t'(s+q)}$$

$$1 + N_{q,s,t'}(i+1) \cdot R_{q,s,t'}(i+1) = 1 + q \cdot N_{q,s,t'}(i+1)^{s+1} =$$

$$= 1 + q \cdot \left[N_{q,s,t'}(i) + 1\right]^{s+1} \leq$$

$$\leq 1 + q \cdot \left[2t' \cdot \left(R_{q,s,t'}(i) + 1\right)^{N_{q,s,t'}(i)}\right]^{s+1} \leq$$

$$\leq 1 + q \cdot \left[2t' \cdot \left(2 R_{q,s,t'}(i)\right)^{N_{q,s,t'}(i)}\right]^{s+1} \leq$$

$$\leq 1 + q \cdot \left[2t' \cdot (2q N_{q,s,t'}(i)^s)^{N_{q,s,t'}(i)}\right]^{s+1} \leq$$

$$\leq 1 + q \cdot \left[(N_{q,s,t'}(i) t' + q + s)^{N_{q,s,t'}(i)}\right]^{s+1} \leq$$

$$\leq 1 + q \cdot \left[2^{(t' + q + s) N_{q,s,t'}(i)^s}\right]^{s+1} \leq$$

$$\leq \exp^{(1)} \left[q + (t' + q + s)(s + 1) \cdot N_{q,s,t'}(i)^2\right] \leq$$

$$\leq \exp^{(1)} \left[2^{(t' + q + s) \cdot N_{q,s,t'}(i)^2}\right] \leq$$

$$\leq \exp^{(1)} \left[2^{(t' + q + s) \cdot \exp^{(i)} [(6 \cdot 2^i - 2) t'(q + s)]}\right] \leq$$

$$\leq \exp^{(2)} \left[2 \exp^{(i)} \left[(6 \cdot 2^i - 2) t'(q + s)\right] + (t' + q + s)\right] \leq$$

$$\leq \exp^{(2)} \left[\exp^{(i)} \left(2 \cdot (6 \cdot 2^i - 2) t'(q + s) + (t' + q + s)\right)\right] \leq$$

$$\leq \exp^{(2)} \left[\exp^{(i)} \left(2 \cdot (6 \cdot 2^i - 2) t'(q + s) + 2 t'(q + s)\right)\right] =$$

$$= \exp^{(2)} \left[\exp^{(i)} \left(6 \cdot 2^i + 2) t'(q + s)\right)\right] =$$

$$= \exp^{(i+2)} \left[(6 \cdot 2^i + 2) t'(q + s)\right]$$
With (3), we then get
\[ |V(T_0)| \leq |V(U_{h,q,s,t})| \leq \prod_{i=0}^{h-1} (1 + R_{q,s,t}(i) \cdot N_{q,s,t}(i)) \leq \prod_{i=0}^{h-1} \exp^{i+1} \left[ (6 \cdot 2^i - 2)q'(q + s) \right] \leq \exp^{(h)} \left[ 2 \cdot (6 \cdot 2^{h-1} - 2)q'(q + s) \right] \leq \exp^{(h)} \left[ (2^{h+3} - 4) \cdot (t + 3q + s)q + s) \right] \leq \exp^{(h)} \left[ (2^{h+5} - 12) \cdot (t + q + s)q + s) \right]. \]

\[ \square \]

**Corollary 3.4.** Let \( T \) be a rooted \( t \)-labelled tree of constant height \( h \geq 1 \), and let \( \phi \) be an MSO sentence with \( r \) quantifiers. Then \( T \models \phi \) can be decided by an FPT algorithm running in elementary time
\[ O \left( \exp^{(h+1)} \left[ 2^{h+5} \cdot r(t + r) \right] + |V(T)| \right) = \exp^{(h+1)} \left( O(|\phi|^2) + O(|V(T)|) \right). \]

**Proof.** Let \( T_0 \) be the kernel obtained in linear time by Theorem 3.2. We (by brute force) exhaustively expand all the quantifiers of \( \phi \) into all possible valuations in \( T_0 \), having at most \( 2^{|V(T_0)|} \) possibilities for each. By searching this “full valuation tree” in time \( O(2^{V(T_0)}(r+1)) \) we decide whether \( T_0 \models \phi \). Using the size bound on \( T_0 \) given by Theorem 3.2, where \( r = q + s \), it is
\[ 2^{V(T_0)}(q+s+1) \leq 2^{\exp^{(h)} \left[ (2^{h+5} - 12) \cdot (t + q + s)q + s) ) \right]} \leq \exp^{(h+1)} \left[ (2^{h+5} - 12) \cdot (t + q + s)q + s) + (q + s + 1) \right] \leq \exp^{(h+1)} \left[ 2^{h+5} \cdot r(t + r) \right]. \]

\[ \square \]

### 3.2 Counting MSO logic

Theorem 3.2 can be further strengthened by considering CMSO logic. Although relatively easy, this is not a simple corollary, and the additional issues of different sort require us to repeat the overall structure of the previous proof as follows.

**Theorem 3.5.** Let \( T \) be a rooted \( t \)-labelled tree of height \( h \), let \( \phi \) be a CMSO sentence with \( q \) element quantifiers and \( s \) set quantifiers, and let \( M \) be the least common multiple of the \( b \) values of all \( \text{mod}_{a,b} \) predicates occurring in \( \phi \).

a) Suppose that \( u \in V(T) \) is a node at level \( i + 1 \) where \( i < h \). If, among all the limbs of \( u \) in \( T \), there are more than \( R_{M+q,s,t+3q,s}(i) \) pairwise \( l \)-isomorphic ones, then let \( T' \subseteq T \) be obtained by deleting exactly \( M \) of the latter limbs from \( T \). Then, \( T \models \phi \iff T' \models \phi. \)
b) Consequently, there is a subtree $T_0 \subseteq T$ computable in linear time (non-parameterized), such that $T \models \phi \iff T_0 \models \phi$ and the size of $T_0$ is bounded by

$$|V(T_0)| \leq \exp(h) \left[ (2^{h+5} - 12) \cdot (t + M + q + s)(M + q + s) \right].$$

Proof. We closely follow the structure of the proof of Theorem 3.2 and implicitly refer to its assumptions and notation. In particular, let $t' = t + 3q + s$. Due to the effects of (future) quantifier elimination onto the $\text{mod}_{a,b}(X)$ predicates we have to deal in $\phi$ also with special model constants $\text{mod}_{a,b}^L$ (where $L$ is a label of the model): The semantics of $\text{mod}_{a,b}^L$ in $U \models \psi$ is that the respective model $U$ contains a modulo $b$ nodes holding label $L$.

a) As in the previous proof, we have a node $u$ in $T$ such that $u$ has many, $p > R_{M+q,s,t'}(i) \geq M$, pairwise $l$-isomorphic limbs $B_1, \ldots, B_p \subseteq T$. We claim that for $T' = T - V(B_1 \cup \cdots \cup B_M)$ it holds $T \models \phi \iff T' \models \phi$. This is again proved by induction on $q + s$:

For the base of induction, when $q = s = 0$, $\phi$ is a propositional formula and the outcomes of $T \models \phi$ and $T' \models \phi$ might differ only in the constants $\text{mod}_{a,b}^L$. However, $M$ is a multiple of $b$ by definition and the number of (deleted) nodes holding label $L$ in $B_1 \cup \cdots \cup B_M$ is a multiple of $M$, and hence all the involved model constants $\text{mod}_{a,b}^L$ in $\phi$ indeed do have the same value over $T$ as over $T'$.

For the induction step with $q + s \geq 1$, we show that the threshold value $p > R_{M+q,s,t'}^L(i) := R_{M+q,s,t'}(i)$ is sufficient in this proof. Note that $R_{M+q,s,t'}^L(i) = M$.

We again proceed by means of contradiction: assuming $T \models \phi$ while $T' \models \neg \phi$ up to symmetry between $\phi$ and $\neg \phi$. Let us consider a counterexample which minimizes the size of $B_1$. Then, analogously to the proof of Theorem 3.2 the $l$-isomorphic limbs $B_1, \ldots, B_p$ are $R_{M+q,s,t'}$-reduced or a smaller counterexample exists (for this $\phi$ or $\neg \phi$). The only difference in the argument is that we are now always removing $M$-tuples of $l$-isomorphic limbs instead of single ones.

Then a quantifier elimination argument, essentially same as previous, finishes the induction step. The argument from the proof of Theorem 3.2 can be simply repeated word by word, only replacing $q$ with $q' := M + q$ as in $R_{q',s,t'}(i)$.

b) By a recursive bottom-up application of a) to the whole tree $T$ we obtain a tree $T_0$ which is $R_{q',s,t'}$-reduced, and this $T_0$ is computable in linear time, too. The size bound then follows from Theorem 3.2 for $q' = M + q$ in place of $q$. □

**Corollary 3.6.** Let $T$ be a rooted $l$-labelled tree of constant height $h \geq 1$, and let $\phi$ be an CMSO sentence with $r$ quantifiers. Let $M$ be the least common multiple of the $b$ values of all $\text{mod}_{a,b}$ predicates occurring in $\phi$. Then $T \models \phi$ can be decided by an FPT algorithm running in elementary time

$$\exp(h+1) \left( O(M + |\phi|)^2 \right) + O(|V(T)|).$$ □

### 4 Algorithmic Consequences for Graphs

If one considers extending algorithmic scope of the previous section to richer classes of structures such as general graphs, the first natural choice would be
to employ efficient interpretation of the structures into coloured trees of fixed height together with Corollaries 3.4 and 3.6. This has been the course taken in the conference version of this paper (covering Lampis [15] and Ganian [11] as special cases) and in a greater generality in [12]. Here we use another, slightly more complicated, approach which has the advantage of being able to smoothly incorporate also some wider problem frameworks, such as MSO enumeration and the LinEMSO optimization framework.

4.1 Smaller tree automaton for MSO

Our approach uses tree automata for MSO properties (cf. Rabin [19]), and its core idea is to give a stricter bound on the number of states of such an automaton by showing that each reachable state is represented by some of our reduced kernels. Recall that the number of states of the automaton related to, e.g., Courcelle’s MSO$_2$ theorem [2] grows non-elementarily with the quantifier alternation depth of the formula, and that this is generally unavoidable by Frick and Grohe [10] already for MSO properties on trees.

For a class of relational structures $S$, we say that an MSO interpretation $I$ of $S$ into a class of rooted trees $T$ is hereditary if, roughly saying, subtrees interpret respective induced substructures. Formally, for any $S \in S$ which is interpreted in $T \in T$ with the domain $D \subseteq V(T)$, every rooted subtree $T' \subseteq T$ interprets by $I$ a structure $S'$ that is the restriction of $S$ onto $D \cap V(T')$. (We implicitly consider simple interpretations.) Our core claim is now formulated as follows.

**Theorem 4.1.** Let $S$ be a hereditary class of binary relational structures, $\psi$ a CMSO property over $S$, and $I$ a simple hereditary CMSO interpretation of $S$ into a class $T$ of rooted (labelled) binary trees. If there exists a simple hereditary CMSO interpretation $J$ of $S$ into the class $U$ of rooted labelled trees of fixed height $h$, then the following holds: There is a finite deterministic tree automaton $A_{\psi,I}$ accepting the language of those $T \in T$ such that $T \models \psi^I$ ($\psi$ under the interpretation $I$); and the number of states of $A_{\psi,I}$ is at most

\[ \exp^{(h+1)} \left( O(M + |\psi^J| + \iota)^2 \right), \]

where (i) $\psi^J$ is $\psi$ under the interpretation $J$, (ii) $M$ is the least common multiple of the $b$ values of all mod$_a$ predicates occurring in $\psi^J$, and (iii) $\iota$ is a constant depending on $I$, $T$ but not on $\psi$.

Recall, again, that sole existence of the automaton $A_{\psi,I}$ follows from Rabin’s [19], while the selling point of this claim is an elementary size bound on $A_{\psi,I}$ for each fixed $h$.

**Proof.** Our approach builds upon the classical Myhill–Nerode regularity tool in automata theory. We actually apply its tree-automata version [14] – we show that the number of classes of the congruence relation on trees defined with respect to $\psi^I$ is bounded as in the statement of the theorem. The existence of the automaton $A_{\psi,I}$ with the same number of states then follows.
For an CMSO sentence $\phi$ we define an equivalence relation $\sim_\phi$ on the universe of rooted (labelled) trees as follows. If $T$ and $B_2$ are rooted trees, and $B_1$ is a limb of a node $v \in V(T)$, then let $T[B_1 \rightarrow B_2]$ denote the tree obtained from $T - V(B_1)$ by attaching the root of $B_2$ as a child of $v$. It is

$$B_1 \sim_\phi B_2 \quad \text{if, and only if,} \quad T \models \phi \iff T[B_1 \rightarrow B_2] \models \phi$$

over all rooted trees $T$ such that $B_1$ is a limb of $T$.

We are interested in the equivalence classes of $\sim_\phi$ when restricted to rooted subtrees of the members of $\mathcal{T}$. Let $S \in \mathcal{S}$ be interpreted by $T \in \mathcal{T}$ (i.e., $S \simeq T^J$), and let $B$ be a limb in $T$. Let $S_1$ and $S_2$ be the induced substructures of $S$ interpreted in $B$ and $T - V(B)$, respectively. Let $U \in \mathcal{U}$ be an interpretation of $S \simeq U^J$ under $J$. Let $U_1, U_2$ be two disjoint copies of $U$, and let $U_3$ be the rooted tree (of height $h + 1$) obtained from $U_1 \cup U_2$ by adding a new root as the parent of the former roots of $U_1, U_2$.

Our first claim is that there exists an assignment of additional labels to $U_3$ and an interpretation $J_3$ of $S$ into $U_3$ such that; (a) $J_3$ depends only on $J$ and on $I, \mathcal{T}$ (this will define our $\iota$), (b) $S_i \simeq U_i^J$ for $i = 1, 2$, and (c) the additional labelling of $U_1$ within $U_3$ is independent of $T - V(B)$.

To prove this claim, consider any of the binary relational symbols $R$ over $S$ and the CMSO formula $\varrho(x, y) \equiv R(x, y)^J$ which interprets $R$ into $\mathcal{T}$. Already by Rabin’s [19]—while simply amended by finite-state mod $\varrho$ predicates in $\varrho$ — this $\varrho(x, y)$ has a fixed finite number of states with the interpretation of $x$ into $V(B) \cap \text{dom}(S)$ and $y$ into $(V(T) \cap \text{dom}(S)) \setminus V(B)$ (and the same applies also to $\varrho(y, x)$ interpreting the inverse $R^{-1}$ in case of a non-symmetric relation).

Hence there exists an integer $m_R$ depending only on $I, \mathcal{T}$ such that $V(S_i)$ has a partition $P_R$ with $|P_R| \leq m_R$ parts, and $S \models R(x_1, y) \iff S \models R(x_2, y)$ for any two $x_1, x_2$ from the same part of $P_R$ and any $y \in V(S_2)$. Importantly, the partition $P_R$ does not depend at all on $S_2$ and $T - V(B)$.

We choose $\iota = 2 + \sum_{R \text{ over } \mathcal{S}} (m_R + m_{R^{-1}})$ and assign $\iota$ new labels to the nodes of $U_3$ as follows: For $i = 1, 2$, a separate new label $L_i$ is issued to the whole domain of $S_i$ in $U_i$. Furthermore, on the domain of $S_1$ in $U_1$, every vertex gets one of $m_R$ new labels identifying which part of $P_R$ it belongs to. This is repeated for all relational symbols $R$ and their inverses $R^{-1}$ over $S$. On the domain of $S_2$ in $U_2$, the label identifying a part $P \in P_R$ is simply given to all the vertices which are adjacent to $P$ via $R$. The interpretation $J_3$ now follows naturally; $U_3 \models R(x, y)^{J_3}$ if, and only if,

\begin{itemize}
  \item $x, y$ in the domain of $S$ are from the same one of $U_1, U_2$—formally, $L_1(x) \land L_1(y)$ or $L_2(x) \land L_2(y)$ —and $U_3 \models R(x, y)^J$ (since $J$ is hereditary), or
  \item $x, y$ in the domain of $S$ are from different ones of $U_1, U_2$—formally, $L_1(x) \land L_2(y)$ or vice versa—and the additional labels in $U_3$ “encode” $R(x, y)$ as defined previously.
\end{itemize}

This finishes the proof of the first claim; the existence of $J_3$. Note, moreover, that $T \models \psi^J \iff S \models \psi \iff U_3 \models \psi^{J_3}$.

14
Second, we claim that every equivalence class of \( \sim_\psi \) contains a special small representative (not necessarily unique), and so the index of \( \sim_\psi \) cannot be too large. We apply Theorem 3.5 a) to \( \psi^{f_3} \) and the tree \( U_3 \), precisely to all nodes of \( U_1 \) within it. Let \( U_1 \) reduce to \( U_0 \) by this application(s), and let \( U_{30} = U_3[U_1 \to U_0] \). Then \( U_3 \models \psi^{f_3} \iff U_{30} \models \psi^{f_3} \). Since \( J \) is hereditary, \( U_{30}^{f_3} \) as a restriction of \( U_0^{f_3} \) is isomorphic to an induced substructure \( S_0 \subseteq S_1 \). Let \( B_0 \subseteq B \) be the subtree giving \( dom(S_0) \) in the interpretation \( I \), and let \( S_{30} = S \setminus (V(S_1) \setminus V(S_0)) \cong U_{30}^{f_3} \). Then, again, \( B_1 \cong S_0 \) and \( T[B \to B_0] \cong S_{30} \) since \( I \) is hereditary. This \( B_0 \) will be a representative of the \( \sim_\psi \)-class of \( B \).

By the previous, we have got \( T \models \psi^I \iff U_{30} \models \psi^{f_3} \iff S_{30} \models \psi \iff T[B \to B_0] \models \psi^I \). This has been so far verified for one particular tree \( T \in \mathcal{T} \).

Consider now arbitrary \( T' \in \mathcal{T} \) having \( B \subseteq T' \) as its limb, and correspondingly define \( S' \in \mathcal{S} \), \( S' \cong T'^I \) and \( U' \in \mathcal{U} \) such that \( S' \cong U'^{f_3} \) by the assumptions. If \( S'_1, S'_2 \subseteq S' \) are the induced substructures interpreted in \( B \) and \( T' - V(B) \), then \( S'_1 = S_1 \). We take \( U'_1 := U_1 \) and \( U'_2 \) a disjoint copy of \( U' \), and analogously construct \( U'_3 \) from \( U'_1 \cup U'_2 \) by adding a new root. As in the first claim, and emphasizing the condition \((c)\), we get a labelling of \( U'_3 \) such that \( S' \cong U'_3^{f_3} \) and \( S'_i \cong U'_i^{f_3} \) for \( i = 1, 2 \). Again by Theorem 3.5 a), for \( U'_{30} = U'_{3}[U'_1 \to U_0] \) it holds \( T' \models \psi^I \iff S' \models \psi \iff U'_3 \models \psi^{f_3} \iff U'_{30} \models \psi^{f_3} \iff T'[B \to B_0] \models \psi^I \). Consequently, \( B \cong_\psi B_0 \) as desired.

Therefore, the number of equivalence classes of \( \sim_\psi \) is at most as large as the number of pairwise non-l-isomorphic \( R_{M+q,s,t+3q+s} \)-reduced rooted \( k \)-labelled trees \( (B_0) \) of height \( h \), where \( q, s \) are the numbers of element and set quantifiers in \( \psi^{f_3} \) and \( t \) is the number of labels addressed in \( \psi^{f_3} \). This quantity is at most \( N_{M+q,s,t+3q+s}(h) \) by Lemma 3.3. Under a very rough estimate, \( q, s \leq t + 3q + s \leq O(|\psi^I| + t) \) by the construction of \( J_3 \). Using the calculation of (4) (for \( i + 1 = h \), and again with a broad margin) we get that \( N_{M+q,s,t+3q+s}(h) < \exp^{(h+1)}(O(M + |\psi^I| + t)^2) \). This is an upper bound on the number of states of desired minimal \( A_{\psi,t} \) by the Myhill–Nerode theorem.

\[ \square \]

### 4.2 Solving extended MSO properties

Unfortunately, direct algorithmic applicability of Corollaries 3.4 and 3.6 is limited to pure decision problems (such as, e.g., 3-colourability), but many practical problems are formulated as optimization ones. The usual way of transforming optimization problems into decision ones does not work for us since the MSO language cannot handle arbitrary numbers.

Nevertheless, there is a known solution. Arnborg, Lagergren, and Seese [1] (while studying graphs of bounded tree-width), and later Courcelle, Makowsky, and Rotics [4] (for graphs of bounded clique-width), specifically extended the expressive power of MSO logic to define so-called LinEMSO optimization problems. Briefly saying, the LinEMSO language allows, in addition to ordinary MSO expressions, to compare between and optimize over linear evaluational terms.

We follow, for an illustration, a simpler definition of LinEMSO\(_1\) given in [4]. Consider any MSO\(_1\) formula \( \psi(X_1, \ldots, X_p) \) with free set variables, and state the
The following problem on an input graph $G$:

$$\text{opt}\{f_{\text{lin}}(U_1, \ldots, U_p) : U_1, \ldots, U_p \subseteq V(G), G \models \psi(U_1, \ldots, U_p)\},$$

where $\text{opt}$ can be $\min$ or $\max$, and $f_{\text{lin}}$ is a linear evaluational function. It is $f_{\text{lin}}(U_1, \ldots, U_p) = \sum_{i=1}^{p} \sum_{j=1}^{m} a_{i,j} \cdot \sum_{x \in U_i} f_j(x)$, where $m$ and $a_{i,j}$ are (integer) constants and $f_j$ are (integer) weight functions on the vertices of $G$. Typically $f_{\text{lin}}$ is just a cardinality function. Such as, $\psi(X) \equiv \forall v, w (v \notin X \lor w \notin X \lor \neg \text{edge}(v, w))$ and “$\max |X|$” describes the maximum independent set problem, or $\psi(X) \equiv \forall v \exists w [v \in X \lor (w \in X \land \text{edge}(v, w))]$ and “$\min |X|$” is the minimum dominating set problem.

The algorithms given in [1,4] for solving such LinEMSO optimization (and enumeration as well) problems are implicitly based on a finite tree automaton associated with the formula $\psi$ in the problem description. Now, Theorem 4.1 can immediately be used to tighten runtime analysis of each of the mentioned algorithms, when the input is restricted to graph classes having not only bounded tree-width or clique-width, respectively, but at the same time being interpretable into a class of trees of fixed height.

This goal first gets us to the following definition:

**Definition 4.2 (Tree-depth [17]).** The closure $\text{cl}(F)$ of a rooted forest $F$ is the graph obtained from $F$ by adding from each node all edges to its descendants. The tree-depth $\text{td}(G)$ of a graph $G$ is one more than the smallest height (distance from the root to all leaves) of a rooted forest $F$ such that $G \subseteq \text{cl}(F)$.

Note that tree-depth is always an upper bound for tree-width. Some useful properties of it can be derived from the following asymptotic characterization: If $L$ is the length of a longest path in a graph $G$, then $\lceil \log_2(L+2) \rceil \leq \text{td}(G) \leq L+1$. See Figure 1 for a simple proof of this, as well as for a more extensive study of tree-depth, we refer the reader to [18 Chapter 6]. Here we need:
Lemma 4.3. Let \( d \) be an integer and \( \mathcal{R}_d \) denote the class of all rooted \( d \)-labelled trees of height \( d \). The class of all graphs of tree-depth at most \( d \) has a simple hereditary MSO\(_1\) interpretation into \( \mathcal{R}_{d+1} \).

Proof. Let \( td(G) \leq d \) and \( W \) be a rooted forest of height \( d \) such that \( G \subseteq cl(W) \), and let \( T \in \mathcal{R}_{d+1} \) be obtained from \( W \) by adding a new common root. The intended interpretation identically maps \( V(G) \) into \( V(T) \). In particular, each vertex quantifier \( \exists x \ldots \) is simply replaced with \( \exists x. \neg L_0(x) \land \ldots \), where \( L_0 \) is a special label given to all vertices in \( V(T) \setminus V(G) \), and nothing is changed with vertex-set quantifiers.

Every vertex \( v \) of \( G \) at distance \( j \geq 1 \) from the root of \( T \) is given the label \( L_j \).

Every vertex \( v \) of \( G \), such that there exists an ancestor \( u \) of \( v \) in \( T \) and \( uv \in E(G) \), is also given the label \( L_i \) where \( i \geq 1 \) is the distance of \( u \) from the root in \( T \).

It is a routine to express the edge relation \( \eta \) of \( G \) as follows:

\[
\eta(x, y) \equiv x \neq y \land (\alpha(x, y) \lor \alpha(y, x)) \text{ where } \alpha(x, y) \equiv \bigvee_{j=1, \ldots, d} (L_j(x) \land L_j(y) \land \bigwedge_{i>j} \neg L_i(x)).
\]

Corollary 4.4. Let \( C \) be a class of graphs of bounded tree-depth. Then every LinEMSO\(_2\) problem \( P \) can be solved on \( C \) by a linear-time FPT algorithm with an elementary runtime dependence on \( P \).

Proof. Let the definition of \( P \) be based on an MSO formula \( \psi \) (for simplicity, \( \psi \) can be thought of an MSO\(_1\) formula thanks to Theorem 2.3). Then, by Lemma 4.3 Theorem 4.1 applies here and the runtime analysis of [1] can be tightened to an elementary bound using (5). \( \square \)

In exactly the same way we can claim an analogous statement for LinEMSO\(_1\):

Corollary 4.5. Let \( C \) be a class of graphs of bounded clique-width, and assume that \( C \) has a simple hereditary MSO\(_1\) interpretation into a class of trees of fixed height. Then every LinEMSO\(_1\) problem \( P \) can be solved on \( C \) by a linear-time FPT algorithm with an elementary runtime dependence on \( P \), provided a clique-width expression for the input graph is given. Then same holds also if counting CMSO\(_1\) is allowed in the description of \( P \). \( \square \)

Remark 4.6. Theorem 4.1 guarantees the existence of a “small” automaton for every CMSO\(_1\) property on classes of graphs interpretable into trees of bounded height, which we use in Corollary 4.5. Unfortunately we cannot directly do the same with CMSO\(_2\) in Corollary 4.4. The reason for this is that Theorem 2.3 only relates MSO\(_1\) and MSO\(_2\) logics, but not CMSO\(_1\) and CMSO\(_2\). However, we believe that an extension with CMSO\(_2\) also holds for Corollary 4.4 in the same way.

In relation to Corollary 4.5 it becomes interesting to ask what graph classes have a MSO\(_1\) interpretation into a class of trees of fixed height. The answer [12] is closely mentioned in other context later in Section 5.2.
5 Consequences for Expressive Power of FO

A non-algorithmic straightforward corollary of Theorem 3.2 is the fact that FO and MSO logic can express the same collection of properties on classes of trees of fixed height; they have the same expressive power. The purpose of this section is to investigate a generalization of this claim to graphs.

5.1 Case of bounded tree-depth

Elberfeld, Grohe, and Tantau [9] proved that FO and MSO have equal expressive power on the graphs of bounded tree-depth. Having Theorem 3.2 at hand, we can provide a relatively simple alternative proof of this result. This alternative proof uses some sophisticated combinatorial tools, namely well-quasi-ordering, which may be interesting on their own in the context of logic. The proof is also a prerequisite for the related new result about FO and MSO in Theorem 5.13.

Theorem 5.1 (Elberfeld, Grohe, and Tantau [9]). Let $D$ denote a class of graphs of bounded tree-depth (Definition 4.2). Then FO and MSO have the same expressive power on $D$.

Proof. Let $\psi_2$ be an MSO sentence. Since MSO has the same expressive power as MSO on graphs of bounded tree-width (by Theorem 2.3), we may as well consider an equivalent MSO sentence $\psi$. Our alternative proof can be then outlined in three steps:

(I) By Lemma 4.3, there is a simple hereditary MSO interpretation $J$ of $D$ into the class $\mathcal{U}$ of rooted labelled trees of height $\leq d$, for some integer constant $d$. By Theorem 3.2, there is a finite set $U_0 \subseteq \mathcal{U}$ of ($f$-reduced) kernels; every $T \in \mathcal{U}$ f-reduces to an “easily definable” $T_0 \in U_0$ such that $T \models \psi^J$ $\iff$ $T_0 \models \psi^J$ (where the reduction threshold $f$ depends only on $\psi, J, d$ and we will shortly say “reduces” in this proof).

(II) For $G \in D$, hence, $G \models \psi$ is equivalent to saying that $G \simeq T^J_G$ for a tree $T^J_G \in \mathcal{U}$ such that $T_G$ reduces to a tree in $\mathcal{U}_0^{\psi^J} = \{ U \in U_0 : U \models \psi^J \}$ (a disjunction of finitely many cases over the members of $U_0^{\psi^J}$).

(III) The problem is that $T^J_G$ is only implicit and we cannot directly address $T^J_G$ (and reducibility of it to $U \in U_0^{\psi^J}$) from within the resulting FO formula over $G$. To resolve this problem, we consider a related hereditary property over $G$ which is characterized by finitely many obstacles by Ding [6]. We use it to build the desired FO sentence expressing over $G$ that (some) implicit $T_G$ reduces to a particular reduced tree $U \in U_0^{\psi^J}$.

The rest of the proof will give the details of crucial step (III).

Let $T \in \mathcal{U}$, and let $\ell(T)$ denote the domain (vertex set) of $T^J$. We say that a graph $H \in D$ is $T$-coloured (with respect to implicit $J$) if $H$ is associated with an injective mapping from $\ell(T)$ into $V(H)$. When dealing with such a coloured graph $H$ we automatically consider a subgraph relation preserving these colours. The key definition is now that of consistency with a $T$-colouring:
For a $T$-coloured graph $G \in \mathcal{D}$, assume there exists an induced supergraph $H \supseteq G$ and a tree $T_H \in \mathcal{U}$ such that $H \simeq T_H^1$. Assume further that $T_H$ is such that $T \subseteq T_H$ (sharing the root with $T$), the isomorphism mapping of $T_H^1 \simeq H$ restricted to $\ell(T)$ is the given $T$-colouring of $G$, and $T_H$ reduces to $T$. In such a situation we say that $G$ is consistent with its $T$-colouring (shortly consistent).

The definition automatically gives that the property of consistency with $T$ is hereditary (closed under induced colour-preserving subgraphs) on the universe of $T$-coloured graphs from $\mathcal{D}$. Therefore, by [15], there is a finite set $Forb(T)$ of $T$-coloured graphs such that $T$-coloured $G$ is consistent if and only if $G$ has no induced subgraph isomorphic to a member of $Forb(T)$. Consequently, this property can be expressed by an FO formula $\text{consistent}_T$.

We also use one special property of the interpretation $J$ from Lemma 4.3 that $J$ is “tree-ordered”, meaning that an edge $vw$ in the graph $T^J$ may exist only if the interpretation of $u$ in $T$ is an ancestor of that of $v$ or vice versa. This property, informally, will allow us to simply identify limbs of $T$ within $T^J$ using connectivity; if $T_1, T_2$ are two disjoint limbs in $T$, then there is no edge between $\ell(T_1)$ and $\ell(T_2)$ in the graph $T^J$.

Now we use the previous to describe the property that $G \in \mathcal{D}$ reduces to $U \in \mathcal{U}_0$, i.e., that there exists $T_G \in \mathcal{U}$ such that $G \simeq T_G^1$ and $T_G$ reduces to $U$. (Recall that if $G$ reduces to $U$, then $G \models \psi \iff U \models \psi^J$.) Considering the restriction of the mapping $T_G^1 \simeq G$ to $\ell(U)$ we say that $G$ reduces to $U$ respecting this $U$-colouring of $G$. We introduce the following shorthand notation:

- For a tree $T \in \mathcal{U}$, let $\vec{x}_T$ denote a collection of variables indexed by the elements of $\ell(T)$, i.e., $\vec{x}_T = (x_v : v \in \ell(T))$. Let $\vec{x}_T \cup \vec{x}_T^v$ stand for a union of two such collections and $z \in \vec{x}_T$ mean that $z = x_v$ for some member $x_v$ of $\vec{x}_T$.
- Let $L(T)$ be the set of all non-$l$-isomorphic limbs $B \subseteq T$ that reach the considered reduction threshold, i.e., if $B$ is of node $v \in V(T)$ and of height $i$, then at least $f(i) - 1$ other limbs of $v$ in $T$ are $l$-isomorphic to $B$. For $B \in L(T)$ we denote by $T \circ B$ a tree such that $T \circ B$ reduces in one step to $T$, i.e., $T \circ B$ results by adding a disjoint copy of $B$ as a sibling of $B$ into $T$.
- We write $G \models \text{consistent}_T(\vec{x}_T)$ in FO to mean that $G$ is consistent with the $T$-colouring given by an assignment of $\vec{x}_T$.
- We write $G \models \text{connected}(\vec{x}_T, y, z)$ to mean that $y, z \notin \vec{x}_T$ and $y$ and $z$ are connected in $G$ by a path avoiding all $\vec{x}_T$. Although connectivity is not in FO in general, our formula connected is FO over $\mathcal{D}$ since graphs of bounded tree-depth are hereditary and have bounded diameter as well [18].

Let $T_1 \in \mathcal{U}$, and let $T_0$ be a limb of $T_1$ or $T_0 = T_1$. We write in FO that $G \models \text{reduce}_{T_1, T_0}(\vec{x}_{T_1}, y)$ with the following intended (and so far informal) meaning: There exists a graph $H$ and a tree $T_H \in \mathcal{U}$ such that $T_H \supseteq T_1$, and

- $G \subseteq H \simeq T_H^1$, the elements of $\ell(T_1)$ coincide with the assignment of $\vec{x}_{T_1}$ in $G$, and $T_H$ reduces to $T_1$,
- the connected component of $G - \vec{x}_{T_1}$ containing $y$ recursively reduces (within the reduction of $T_H$ onto $T_1$) to $T_0$.

19
This is formally written down as follows

\[
\text{reduce}_{T_1,T_0}(x_{T_1}, y) \equiv \bigvee_{T_2 \in \mathcal{L}(T_0)} \exists \tilde{z}_{T_2}. \ \Psi_{T_1,T_0,T_2}
\]

where

\[
\Psi_{T_1,T_0,T_2} \equiv \bigwedge_{v \in V(T_2)} \text{connected}(x_{T_1}, y, z_v) \wedge \text{consistent}_{T_1 \cup T_2}(x_{T_1} \cup \tilde{z}_{T_2}) \wedge \\
\forall y' \left[\left(\text{connected}(x_{T_1}, y, y') \wedge y' \notin \tilde{z}_{T_2}\right) \rightarrow \text{reduce}_{T_1 \cup T_2}(x_{T_1} \cup \tilde{z}_{T_2}, y')\right].
\]

except that for \(T_2\) of height 0 (i.e., \(T_2\) a single node, hence singleton \(\tilde{z}_{T_2}\)) it is

\[
\Psi_{T_1,T_0,T_2} \equiv y \in \tilde{z}_{T_2} \wedge \forall y'(\text{connected}(x_{T_1}, y, y') \rightarrow y' = y) \wedge \\
\text{consistent}_{T_1 \cup T_2}(x_{T_1} \cup \tilde{z}_{T_2}).
\]

Note that in (6) the only role of \(y\) is to be a representative of a connected component of \(G - x_{T_1}\); the outcome of \(G \models \text{reduce}_{T_1,T_0}(x_{T_1}, y)\) is invariant upon the choice of \(y\) from the same component. Task (III) is now finished with

\[
\text{reduce-to}_U \equiv \exists \tilde{x}_U \left\{ \text{consistent}_U(\tilde{x}_U) \wedge \forall z \left( z \in \tilde{x}_U \lor \text{reduce}_{U,U}(\tilde{x}_U; z) \right) \right\}
\]

Assume that \(G \in \mathcal{D}\) reduces to \(U \in \mathcal{U}_0\). Then, following the steps of this reduction, it is a simple routine to verify that \(G \models \text{reduce-to}_U\) (one always chooses \(H = G\) when speaking about consistency). On the other hand, assume \(G \models \text{reduce-to}_U\). Fix a satisfying assignment of \(\tilde{x}_U\) into \(V(G)\) as the \(U\)-colouring of \(G\). If \(|\ell(U)| = |V(G)|\), then \(G \cong U\) by \(G \models \text{consistent}_U(\tilde{x}_U)\). Otherwise, for every connected component \(P_0\) of \(G - \tilde{x}_U\) choose \(p \in P_0\); hence \(G \models \text{reduce}_{U,U}(\tilde{x}_U; p)\). We now analyze the recursive definition (6) of \(\text{reduce}_{U,U}\), aiming to show the following claim: the induced subgraph \(G_1 = G[\tilde{x}_U \cup P_0]\) reduces to \(U\) respecting the \(U\)-colouring induced by \(\tilde{x}_U\). If this is true for each component \(P_0\) then, clearly, \(G\) itself reduces to \(U\).

Let initially \(U_0 = W_0 = U\), and set \(W_1 = T_2\) where \(T_2 \in \mathcal{L}(W_0)\) of height \(> 0\) is a satisfying branch (any one) of the “big” disjunction in (6) and fix a satisfying assignment of \(\tilde{z}_{W_1}\) into \(P_0\) thereafter. Let \(U_1 = U_0 \cap W_1\) and (with a negligible abuse of notation) \(\tilde{x}_{U_1} = \tilde{x}_{U_0} \cup \tilde{z}_{W_1}\). If \(|\ell(W_1)| = |P_0|\), then we are immediately done with the claim by \(\text{consistent}_{U_1}(\tilde{x}_{U_1})\) in (7): according to the hereditary property of consistency it is \(G_1 \cong U_1\) and ‘\(U_1 \rightarrow U_0\)’ is a valid reduction step. Otherwise, recursively for every connected component \(P_1\) of \(G_1 - \tilde{x}_{U_1}\) (which is one of components of \(G - \tilde{x}_{U_1}\), too) we choose \(q \in P_1\); hence (8) \(G \models \text{reduce}_{U_1,W_1}(\tilde{x}_{U_1}; q)\).

In each recursion branch (say, by DFS on the recursion tree) we continue with the previous argument at “depth” \(i \geq 1\); we analyze (6) \(G \models \text{reduce}_{U_i,W_i}(\tilde{x}_{U_i}; q)\) over every connected component \(P_i \ni q\) of \(G_i - \tilde{x}_{U_i}\), denote by \(G_{i+1} = G[\tilde{x}_{U_i} \cup P_i]\),
and analogously obtain $W_{i+1} (= T_2 \in \mathcal{L}(W_i))$ from a satisfying branch) and $U_{i+1} = U_i \odot W_{i+1}$, with an assignment of $\overline{z_{W_{i+1}}}$ into $P_i$. This recursion is finite since the height of $W_{i+1}$ strictly decreases and then generally terminates at (9) for $W_{i+1}$ of height 0. In this terminal case we get by (9) for $\Psi_{U_i, W_i, W_{i+1}}$ that the assignment of $\overline{z_{W_{i+1}}}$ is $q$ and $\{q\} = P_i$, and so $U_{i+1} \simeq G_{i+1}$ by consistent $U_{i+1}$. This finishes the proof since $U_{i+1}$ by the construction reduces to $U$. □

Notice that much of the previous proof is very general, using just the fact that the graph class $\mathcal{D}$ has some simple hereditary MSO$_1$ interpretation $J$ into the class $\mathcal{U}$. Interestingly, well-quasi-ordering of $\mathcal{D}$ under induced subgraphs can be derived already from this assumption and [6]. There is, however, one technical point in the proof which heavily depends on special properties of the interpretation $J$—it is the use of predicate $\text{connected}(\overline{x_T}, y, z)$ in the proof of Theorem 5.1. This is based on a quite unique property of tree-depth, and before attempting to give a similar result for other graph classes interpretable into $\mathcal{U}$, we have to find a suitable replacement property in the next section.

### 5.2 Tree-models and their properties

Graph classes of bounded shrub-depth have been introduced recently in [12] as those having simple MSO$_1$ interpretations into the class(es) of rooted labelled trees of fixed height. Equivalently, they are defined by a very special kind of a simple hereditary interpretation:

**Definition 5.2 (Tree-model [12]).** We say that a graph $G$ has a tree-model 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 ($T$ is $m$-coloured),
- iv. and the existence of a $G$-edge between $u, v \in V(G)$ depends solely on 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 $\mathcal{T}\mathcal{M}_m(d)$.

For example, $K_n \in \mathcal{T}\mathcal{M}_1(1)$ or $K_{n,n} \in \mathcal{T}\mathcal{M}_2(1)$. Definition 5.2 is further illustrated in Figure 2. It is easy to see that each class $\mathcal{T}\mathcal{M}_m(d)$ is closed under complements and induced subgraphs, but neither under disjoint unions, nor under subgraphs.
Definition 5.3 (Shrub-depth [12]). A class of graphs \( S \) has shrub-depth \( d \) if there exists \( m \) such that \( S \subseteq TM_m(d) \), while for all natural \( m \) it is \( S \not\subseteq TM_m(d-1) \).

Note that Definition 5.3 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). For instance, the class of all cliques has shrub-depth 1. Similarly, although Definition 5.2 does not explicitly specify rules for the existence of edges (iv.), Definition 5.3 provides a natural associated FO interpretation \( J \) of the class \( S \):

Definition 5.4 (Shrub interpretation). For a graph class \( S \) of shrub-depth \( d \), a shrub interpretation of \( S \) into the class \( U \) of rooted labelled trees of height \( d \) is a simple hereditary interpretation \( J \) satisfying the following: for \( G \in S \) with a tree-model \( T \) of depth \( d \), it is \( G \simeq T_1^J \) where \( T_1 \) inherits \( T \) with all the leaf colours, and each leaf \( v \) of \( T_1 \) is additionally equipped with all labels of the form \((i,c)\) where \( c \) is the \( T \)-colour of a vertex \( u \) adjacent to \( v \) in \( G \) such that the distance between \( u,v \) in \( T \) is \( 2i \).

Lemma 5.5. A shrub interpretation is FO definable for each fixed \( d \).  

For more relations of shrub-depth to other established concepts such as cographs or clique-width we refer the reader to [12]. Here we just summarize:

Proposition 5.6 ([12]). Let \( S \) be a graph class and \( d \) an integer. Then:

a) If \( S \) is of tree-depth \( \leq d \), then \( S \) is of shrub-depth \( \leq d \).
b) If \( S \) is of bounded shrub-depth, then \( S \) is of bounded clique-width.

Proposition 5.7 (Ding [6], also [12] in greater generality). Let \( S \) be a graph class of bounded shrub-depth. Assume the graphs of \( S \) are arbitrarily coloured from a finite set of colours. Then \( S \) is well-quasi-ordered under the colour-preserving induced subgraph order.

Proof. Consider an infinite sequence \((G_1,G_2,\ldots)\subseteq S\), and the corresponding tree-models \((T_1,T_2,\ldots)\). Let \( T_i^+, \ i = 1,2,\ldots \), denote the rooted tree with leaf labels composed of the colours of \( T_i \) and the colours in \( G_i \). By Ding [6] (equivalently, by a \( d \)-fold application of folklore Higman’s lemma), \( T_1^+, T_2^+, \ldots \) of bounded height are WQO under rooted coloured subtrees, and, consequently, so are the coloured graphs \( G_1, G_2, \ldots \), as desired.

For the rest of this section we will focus on the shrub interpretation \( J \) associated with a graph class \( S \) of bounded shrub-depth. This allows us to smoothly adopt the notions introduced during the proof of Theorem 5.1. In particular, \( \ell(T) \) (the domain of \( T^J \)) is now the set of leaves of \( T \) and the notion of a \( T \)-colouring corresponds to that. We also literally adopt the definition of \( G \in S \) being consistent with its \( T \)-colouring. Thanks to Proposition 5.7, this property can be expressed by an FO formula \( \text{consistent}_T \) (depending on \( S \)), too.
Though, we need a bit stronger and crucial property of separability for $T$. It deals with a $(T, red, blue)$-coloured graph which is a $T$-coloured graph in which two other (non-$T$-coloured) vertices are assigned colours red and blue. When the red vertex is $x$ and the blue one is $y$ in a graph $H$, then we call this also a $(T, x, y)$-colouring of $H$.

**Definition 5.8 (Separability for $(T, red, blue)$-coloured).** Let $S$ be a class of bounded shrub-depth with a shrub interpretation $J$ into $U$ (cf. Definition 5.4). Assume a $(T, red, blue)$-coloured graph $G$ where $r, b$ denote the vertices of colours red and blue, respectively. We say that $G$ is separating ($r$ from $b$, implicitly) for $T$ if there exists an induced supergraph $H \supseteq G$ and a tree $T_H \in U$, $H \cong T_H'$, such that the following hold:

- $T \subseteq T_H$ (sharing the root with $T$), the isomorphism mapping of $T_H'$ restricts to $H$ to the given $T$-colouring of $G$, and $T_H$ reduces to $T$;
- the least common ancestor of the nodes interpreting $r, b$ within $T_H$ belongs to $V(T)$.

Note: separability clearly implies consistency with $T$. And again by Proposition 5.7, the separability property can be expressed with an FO formula $\text{separable}_T(x_T, r, b)$ over the universe of $(T, r, b)$-coloured graphs from $S$.

To make practical use of the (generally vague) separability property, we have to restrict our domain to so called unsplittable tree-models, as follows.

Assume $T$ is a tree-model of a graph $G$, and $B$ is a limb of $v \in V(T)$, such that $W$ is the set of leaves of $B$. We say that a tree-model $T'$ is obtained from $T$ by splitting $B$ along $X \subseteq W$ if a disjoint copy $B'$ of $B$ with the same parent $v$ is added into $T$, and then $B$ is restricted to the Steiner tree of $W \setminus X$ while $B'$ is restricted to the Steiner tree of $X'$ (the corresponding copy of $X$). A tree-model $T$ is splittable if some limb in $T$ can be split along some subset $X$, making a tree-model $T'$ which represents the same graph $G$ as $T$ does. A tree-model is unsplittable if it is not splittable. Notice that any tree-model can be turned into an unsplittable one; simply since the splitting process must end eventually.

An unsplittable shrub interpretation $J$ is an interpretation satisfying Definition 5.4 with the target class $U$ containing only trees of unsplittable tree-models. We also implicitly assume that the threshold function $f$ in the definition of ‘$f$-reduce’ is always at least 2. Then we claim:

**Lemma 5.9.** Assume an unsplittable shrub interpretation $J$ into $U$. For a graph $G \simeq T_G'$, let $T_G \in U$ reduce to $T_0$, and associate $G$ with this $T_0$-colouring. Let $L_1, \ldots, L_k$ denote the leaf sets of all $k$ subtrees in $T_G \setminus V(T_0)$. If $\sim$ is the binary relation on $V(G) \setminus \ell(T_0)$ defined as $x \sim y$ iff the corresponding $(T_0, x, y)$-colouring of $G$ is not separating, then $\sim$ is a relation of equivalence whose classes coincide with the sets $L_1, \ldots, L_k$.

Before giving a proof, we add the following immediate corollary which may be interesting on its own (though not used here).
Corollary 5.10. If $J$ is unsplittable, then the partition $L_1,\ldots,L_k$ from Lemma 5.9 is unique for given $G$ and $T_0$ (regardless of $T_G$).

Remark 5.11. Note that it is not possible to simply relate which equivalence class of $\sim$ corresponds to which subtree of $T_G - V(T_0)$. This is since the definition of separability deals with a supergraph $H \supseteq G$ and then a tree $T_H \supseteq T_G$ may reduce to $T_0$ in a different way than $T_G$ does.

Proof (Lemma 5.9). Suppose that $x, y$ belong to distinct ones of the sets $L_1,\ldots,L_k$. Then already $H = G$ in Definition 5.8 witnesses that the $(T_0, x, y)$-coloured graph $G$ is separating, and so $x \not\sim y$.

In the other direction, we aim for a contradiction. Two vertices $p, q$ are called twins with respect to a set $R$ if the neighbourhood of $p$ in $R$ equals that of $q$ in $R$. Moreover, $P, Q \subseteq V(G)$ are twin sets in $G$ if there is a bijection $h : P \to Q$ such that $h$ is an isomorphism of $G[P]$ onto $G[Q]$, and $p, h(p)$ are twins wrt. $V(G) \setminus (P \cup Q)$ for each $p \in P$.

We assume (up to symmetry) that $x, y \in L_1$, but the $(T_0, x, y)$-coloured graph $G$ is separating for $T_0$. Hence there is a graph $H \supseteq i$, and a tree $T_H \supseteq T_0$, $H \cong T_H^i$, such that the least common ancestor of $x, y$ within $T_H$ belongs to $V(T_0)$. If $K_1,\ldots,K_{k'}$ denote the sets of leaves of all $k'$ subtrees in $T_H - V(T_0)$ then, say, $x \in K_1$ but $y \not\in K_1$. We summarize what this means according to our definitions:

i. Since $T_G$ reduces to $T_0$, the set $L_1$ is the leaf set of a subtree $U_1$ in $T_G - V(T_0)$, and this $U_1$ is reducible to $U_1'$ such that $U_1'$ is isomorphic to (at least) two disjoint sibling limbs $B, B'$ in $T_0$. Clearly, we may assume that $x, y$ are in $U_1'$. We consider the induced subgraph $G' = G - (V(U_1) \setminus V(U_1'))$, and $L_1', M, M' \subseteq V(G')$ the leaf sets of $U_1', B, B'$, respectively. Then $L_1', M, M'$ are pairwise twin sets in $G'$ by Definition 5.2.

ii. The same as in (i) can be claimed for $H, T_H, and K_1$ in place of $G, T_H$, and $L_1$; giving us sets (pairwise twin) $K_1' \subseteq K_1$ and $N, N' \subseteq \ell(T_0)$ in a suitable induced subgraph of $H$. We may similarly assume $x \in K_1'$. Moreover, since $N, N'$ are the leaf sets of some limbs in $T_0$, too, each of $N, N'$ may either be disjoint or in an inclusion with each of $M, M'$. Consequently, up to symmetry, $N$ can be assumed disjoint from $M'$. In the graph $G'$ we, in particular, have that for every vertex $u \in K_1' \cap V(G')$ there is $w \in N$ such that $u, w$ are twins wrt. $(L_1' \cup M \cup M') \setminus (K_1' \cup N)$.

Now, we get a contradiction (to being unsplittable) by applying next Lemma 5.12 with $G := G', W := M, W' := M', Y := L_1', X := N, Z := K_1' \cap V(G')$.

Lemma 5.12. Let $G$ be a graph, and $T_0$ be a tree-model of an induced subgraph $G_0 \subseteq i$. Let $T_0$ contain two (disjoint) isomorphic limbs $B, B'$ of a node $v$, and $W, W' \subseteq V(G_0)$ be the sets of leaves of $B, B'$, respectively. Let $X \subseteq V(G_0)$ be a set disjoint from $W'$. If there exist sets $Y, Z \subseteq V(G) \setminus V(G_0)$ such that

a) $W, W', Y$ are pairwise twin sets in $G_0$,
Fig. 3. A situation which cannot happen, in a graph $G$ with an unsplittable tree-model $T_0$ of an induced subgraph $G_0 \subseteq G$, and with the sets $W, W', Y$ and $X, Z$ as in Lemma 5.12.

b) for each $z \in Z$ there is some $h(z) \in X$ such that $z, h(z)$ are twins with respect to $(W \cup W' \cup Y) \setminus (Z \cup X)$,

c) $Y \neq Y \cap Z \neq \emptyset$,

then the tree model $T_0$ is splittable.

The statement is illustrated in Figure 3.

Proof. Our aim is to prove that, assuming a) – c) to be true, the limb $B$ in $T_0$ can be split along the set $W \subseteq W$ which in the twin set relation (a) corresponds to $Y \cap Z$ (and analogously for $B'$). We take arbitrary $z \in Y \cap Z$ and $y \in Y \setminus Z$, and denote by $w, t \in W$ and $w', t' \in W'$ the vertices such that $w, w'$ correspond to $y$ and $t, t'$ to $z$ in the twin set relation (in particular, $t \in W_Z$). By Definition 5.2, only the edges between $W_Z$ and $W \setminus W_Z$ in $G_0$ are potentially affected by the splitting operation on $B$, and so it is enough to show that $w \in E(G_0)$ iff $w \in E(G_0)$ to get the desired conclusion.

Assume $w \in E(G_0)$. Then $yz \in E(G)$ by (a) for the twin pair $W, Y$. Since $y \notin Z \cup X$, from (b) and $yz \in E(G)$ we get $yh(z) \in E(G)$. Then, $h(z) \in X$ implies $h(z) \notin W'$ which is disjoint from $X$, and so from (a) for the twin pair $W', Y$ it follows $w'h(z) \in E(G)$. Furthermore, $w' \notin Z \cup X$, and hence $w'z \in E(G)$ from (b). Finally, $w't \in E(G)$ by (a) for the twin pair $W, Y$, and $w't \in E(G_0)$. In the exactly same way, $w \notin E(G_0)$ implies $w't \in E(G_0)$.

5.3 Case of bounded shrub-depth

We get to the main new result of Section 5.

Theorem 5.13. Let $\mathcal{S}$ denote any class of graphs of bounded shrub-depth. Then $\text{FO}$ and $\text{MSO}_1$ have the same expressive power on $\mathcal{S}$.

Proof. Our proof follows the same steps as that of Theorem 5.1.
(I) Let $J$ be a shrub interpretation (Lemma 5.5) of $S$ into the class $U$ of rooted labelled trees of height $\leq d$, for some integer constant $d$. By implicitly restricting $U$ we may assume it is unsplittable. By Theorem 3.2, there is a finite set $U_0 \subseteq U$ of reduced kernels; every $T \in U$ reduces to an “easily definable” $T_0 \in U_0$ such that $T \models \psi^J \iff T_0 \models \psi^J$.

(II) For $G \in S$, hence, $G \models \psi$ is equivalent to saying that $G \simeq T_G^J$ for a tree $T_G \in U$ such that $T_G$ reduces to a tree in $\{ U \in U_0 : U \models \psi^J \}$.

(III) It remains to build a desired FO sentence expressing over $G$ that (some) implicit unsplittable $T_G$ reduces to a particular unsplittable tree $U \in U_0^\psi,J$.

The rest of the proof will again give the details of crucial step (III). Using the tools developed in Section 5.2 this is now a relatively easy task:

$$reduce_{T_1,T_2}(\widehat{x_{T_1}}, y) \equiv \text{sep-is-equivalence}_{T_1}(\widehat{x_{T_1}}) \land \bigvee_{T_2 \in \mathcal{L}(T_2)} \exists \widehat{z_{T_2}}. \Psi_{T_1,T_0,T_2}$$

where

$$\Psi_{T_1,T_0,T_2} \equiv \bigwedge_{v \in V(T_2)} \neg \text{separable}_{T_1}(\widehat{x_{T_1}}, y, z_v) \land \text{consistent}_{T_1 \circ T_2}(\widehat{x_{T_1}} \cup \widehat{z_{T_2}}) \land \forall y' \left[ (\neg \text{separable}_{T_1}(\widehat{x_{T_1}}, y, y') \land y' \notin \widehat{z_{T_2}}) \rightarrow reduce_{T_1 \circ T_2,T_2}(\widehat{x_{T_1}} \cup \widehat{z_{T_2}}, y') \right] ,$$

except that for $T_2$ of height 0 (i.e., $T_2$ a single node, hence singleton $\widehat{z_{T_2}}$) it is

$$\Psi_{T_1,T_0,T_2} \equiv y \in \widehat{z_{T_2}} \land \forall y' (\neg \text{separable}_{T_1}(\widehat{x_{T_1}}, y, y') \rightarrow y' = y) \land \text{consistent}_{T_1 \circ T_2}(\widehat{x_{T_1}} \cup \widehat{z_{T_2}}).$$

Here ‘sep-is-equivalence$_T$($\widehat{x_{T}}$)’ asserts that the binary relation defined by ‘$\text{separable}_T$($\widehat{x_{T}}, \cdot$)’ is reflexive, symmetric, and transitive on $V(G) \setminus \widehat{x_{T}}$, which has a routine FO expression. Task (III) is now finished with

$$\text{reduce-to}_U \equiv \exists \widehat{x_U} \left\{ \text{consistent}_U(\widehat{x_U}) \land \forall z (z \in \widehat{x_U} \lor \text{reduce}_{U,U}(\widehat{x_U}, z)) \right\} .$$

We finally claim that $G \in S$ reduces to $U \in U_0$ if, and only if, $G \models \text{reduce-to}_U$. Under Lemma 5.9 this is just a plain repetition of the arguments in the proof of Theorem 5.1. $\Box$

6 Conclusions

Despite our prime motivation was to provide an algorithmic improvement over Courcelle’s theorem on “low-depth” variants of tree-width and clique-width, the main importance of our results probably lies in the fact that they provide deeper
understanding of MSO logic on graph classes interpretable in trees of bounded height. This understanding already played crucial role in the proofs establishing that FO logic has the same expressive power as MSO on these graph classes.

We believe that our results and techniques for obtaining them (especially the use of well-quasi-ordering) could lead to new results in the future. Above all we suggest that there is likely no major obstacle to suitable extensions of the results of Sections 4 and 5 to classes of general relational structures. In particular, the notion of shrub-depth extends easily there.

We conclude with a conjecture stating the converse of our result on expressive power or FO and MSO on graph classes of bounded shrub-depth.

Conjecture 6.1. Consider a hereditary (i.e., closed under induced subgraphs) graph class $\mathcal{G}$. If the expressive powers of FO and $\text{MSO}_1$ are equal on $\mathcal{G}$, then the shrub-depth of $\mathcal{G}$ is bounded (by a suitable constant).

Resolving this conjecture would probably require an asymptotic characterization of shrub-depth in terms of forbidden induced subgraphs. Such characterization would probably also allow us to answer the following question: Are there hereditary graph classes of unbounded shrub-depth having an $\text{MSO}_1$ model-checking algorithm with elementary run-time dependence on the formula? For tree-depth and $\text{MSO}_2$ the answer is no (unless EXP=\text{NEXP}), because unbounded tree-depth implies the existence of long paths and by the result of Lampis [16] this in turn implies non-existence of such an algorithm.

References

1. S. Arnborg, J. Lagergren, and D. Seese. Easy problems for tree-decomposable graphs. *J. Algorithms*, 12(2):308–340, 1991.
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. The monadic second order logic of graphs VI: on several representations of graphs by relational structures. *Discrete Applied Mathematics*, 54(2-3):117–149, 1994.
4. 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.
5. R. Diestel. *Graph Theory*, volume 173 of *Graduate texts in mathematics*. Springer, New York, 2005.
6. G. Ding. Subgraphs and well-quasi-ordering. *Journal of Graph Theory*, 16(5):489–502, 1992.
7. R. Downey and M. Fellows. *Parameterized complexity*. Monographs in Computer Science. Springer, 1999.
8. Z. Dvořák, D. Kráľ', and R. Thomas. Deciding first-order properties for sparse graphs. In *FOCS*, pages 133–142, 2010.
9. M. Elberfeld, M. Grohe, and T. Tantau. Where first-order and monadic second-order logic coincide. In *LICS*, pages 265–274, 2012.
10. 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.
11. R. Ganian. Twin-cover: Beyond vertex cover in parameterized algorithmics. In *IPEC’11*, volume 7112 of *LNCS*, pages 259–271. Springer, 2012.
12. R. Ganian, P. Hliněný, J. Nešetřil, J. Obdržálek, P. O. de Mendez, and R. Ramadurai. When trees grow low: Shrubs and fast MSO1. In B. Rovan, V. Sassone, and P. Widmayer, editors, *MFCS*, volume 7464 of *Lecture Notes in Computer Science*, pages 419–430. Springer, 2012.
13. W. Hodges. *A Shorter Model Theory*. Cambridge University Press, New York, NY, USA, 1997.
14. D. Kozen. On the Myhill-Nerode theorem for trees. *Bulletin of the EATCS*, 47:170–173, 1992.
15. M. Lampis. Algorithmic meta-theorems for restrictions of treewidth. *Algorithmica*, 64(1):19–37, Sept. 2012.
16. M. Lampis. Model checking lower bounds for simple graphs. In F. V. Fomin, R. Freivalds, M. Z. Kwiatkowska, and D. Peleg, editors, *ICALP (1)*, volume 7965 of *Lecture Notes in Computer Science*, pages 673–683. Springer, 2013.
17. J. Nešetřil and P. Ossona de Mendez. Tree-depth, subgraph coloring and homomorphism bounds. *European J. Combin.*, 27(6):1024–1041, 2006.
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. Decidability of second-order theories and automata on infinite trees. *Transactions of the American Mathematical Society*, 141:1–35, July 1969.