





Distributed Model Checking on Graphs of Bounded Treedepth

Fedor V. Fomin  

University of Bergen, Norway

Pierre Fraigniaud  

IRIF, Université Paris Cité and CNRS, Paris, France

Pedro Montealegre  

Universidad Adolfo Ibañez, Santiago, Chile

Ivan Rapaport  

Universidad de Chile, Santiago, Chile

Ioan Todinca  

LIFO, Université d'Orléans, France

INSA Centre-Val de Loire, Orléans, France

Abstract

We establish that every monadic second-order logic (MSO) formula on graphs with bounded treedepth is decidable in a constant number of rounds within the CONGEST model. To our knowledge, this marks the first meta-theorem regarding distributed model-checking. Various optimization problems on graphs are expressible in MSO. Examples include determining whether a graph G has a clique of size k , whether it admits a coloring with k colors, whether it contains a graph H as a subgraph or minor, or whether terminal vertices in G could be connected via vertex-disjoint paths. Our meta-theorem significantly enhances the work of Bousquet et al. [PODC 2022], which was focused on *distributed certification* of MSO on graphs with bounded treedepth. Moreover, our results can be extended to solving optimization and counting problems expressible in MSO, in graphs of bounded treedepth.

2012 ACM Subject Classification Theory of computation → Distributed algorithms; Theory of computation → Verification by model checking

Keywords and phrases proof-labeling schemes, local computing, CONGEST model

Digital Object Identifier 10.4230/LIPIcs.DISC.2024.25

Related Version *Full Version*: <https://arxiv.org/abs/2405.03321> [28]

Funding *Fedor V. Fomin*: the research leading to these results has received funding from the Research Council of Norway via the project BWCA (grant no. 314528).

Pierre Fraigniaud: partially supported by the ANR projects DUCAT (ANR-20-CE48-0006), and QuData (ANR-18-CE47-0010).

Pedro Montealegre: P.M. has received funding from FONDECYT 1230599.

Ivan Rapaport: I.R. has received funding from FB210005, BASAL funds for centers of excellence from ANID- Chile, and FONDECYT 1220142.

1 Introduction

Distributed *decision* [12, 30, 31] and distributed *certification* [32, 43, 51] are two complementary fields of distributed computing, closely associated with distributed fault-tolerant computing [3, 19, 50]. Both fields are addressing the problem of checking whether a distributed system is in a legal state with respect to a given specification, or not. We examine this problem in the classical context of distributed computing in networks, under the standard



© Fedor V. Fomin, Pierre Fraigniaud, Pedro Montealegre, Ivan Rapaport, and Ioan Todinca; licensed under Creative Commons License CC-BY 4.0

38th International Symposium on Distributed Computing (DISC 2024).

Editor: Dan Alistarh; Article No. 25; pp. 25:1–25:20



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

CONGEST model [60]. Recall that this model assumes networks modeled as simple connected n -node graphs, in which every node is provided with an identifier on $O(\log n)$ bits that is unique in the network. Computation proceeds synchronously as a sequence of *rounds*. At each round, every node sends a message to each of its neighbors in the graph, receives the messages sent by its neighbors, and performs some individual computation. A crucial point is that messages are restricted to be of size $O(\log n)$ bits. While this suffices to transmit an identifier, or a constant number of identifiers, transmitting large messages may require multiple rounds, typically $\Theta(k/\log n)$ rounds for k -bit messages.

Distributed Decision

Given a boolean predicate Π on graphs, e.g., whether the graph is H -free for some fixed graph H , a *decision* algorithm for Π takes as input a graph $G = (V, E)$, and outputs whether G satisfies Π or not. Specifically, every node v receives as input its identifier $\text{id}(v)$, and, after a certain number of rounds of communication with its neighbors, it outputs *accept* or *reject*, under the constraint that G satisfies Π if and only if the output of each of the nodes $v \in V$ is *accept*. In other words,

$$G \models \Pi \iff \forall v \in V(G), \text{output}(v) = \text{accept}.$$

Some predicates are easy to decide *locally*, i.e., in a constant number of rounds. A canonical example is checking whether the (connected) graph G is regular, for which one round suffices. However, other predicates cannot be checked locally, with canonical example checking whether there is a unique node of degree 3 in the network. Indeed, checking this property requires $\Omega(n)$ rounds in networks of diameter $\Theta(n)$, as two nodes of degree 3 may be at arbitrarily large distances in the graph. Another example of a difficult problem is checking whether the graph is C_4 -free, i.e., does not contain a 4-cycle as a subgraph, which requires $\tilde{\Omega}(\sqrt{n})$ rounds [13]. One way to circumvent the difficulty of local checkability, i.e., to address graph predicates requiring a large number of rounds for being decided, is to consider distributed *certification*.

Distributed Certification

A *certification scheme* for a boolean predicate Π is a pair *prover-verifier* (see [19] for more details). The prover is a centralized, computationally unbounded, non-trustable oracle. Given a graph $G = (V, E)$, the prover assigns a *certificate* $c(v) \in \{0, 1\}^*$ to each node $v \in V$. These certificates are forged by the prover using the complete knowledge of the graph G . The verifier is a distributed 1-round algorithm. Each node v takes as sole input its identifier $\text{id}(v)$ and its certificate $c(v)$. In particular, for distributed decisions, v is unaware of the graph G . Every node v just exchanges once its identifier and certificate with its neighbors, and then it must output *accept* or *reject*.

The certification scheme is correct if the following two conditions hold. The *completeness* condition states that if G satisfies Π , then the oracle can provide the nodes with certificates that they all accept. The *soundness* condition says that if G does not satisfy Π , then no matter the certificates assigned by the oracle to the nodes, at least one of them rejects. That is, the role of the verifier is to check that the collection of certificates assigned to the nodes by the prover is indeed a global proof that the graph satisfies the predicate. In other words,

$$G \models \Pi \iff \exists c : V(G) \rightarrow \{0, 1\}^* : \forall v \in V(G), \text{output}(v) = \text{accept}.$$

The main measure of complexity of a certification scheme is the maximum *size* of the certificates assigned by the prover to the nodes on legal instances, i.e., for graphs G satisfying the given predicate. Ideally, to be implemented in a single round under the CONGEST model, the certificates should be of size $O(\log n)$ bits. Interestingly, many graph properties can be certified with such short certificates, including acyclicity [51], planarity [25], bounded genus [17, 26], etc. On the other hand, basic graph properties require large certificates, including diameter 2 vs. 3 (requiring $\tilde{\Omega}(n)$ -bit certificates [8]), non-3-colorability (requiring $\tilde{\Omega}(n^2)$ -bit certificates [43]), C_4 -freeness (requiring $\tilde{\Omega}(\sqrt{n})$ -bit certificates [13]), etc. The following question was thus raised, under different formulations (see, e.g., [20]): *What are the graph properties that admit certification schemes with $O(\log n)$ -bit certificates, or, to the least, certificates of polylogarithmic size?* Answering this question requires formalizing the notion of “graph predicate”.

Monadic Second-Order Logic

Recall that, in the *first-order* logic (FO) of graphs, a graph property is expressed as a quantified logical sentence whose variables represent vertices, with predicates for equality (=) and adjacency (adj). An FO formula is therefore constructed according to the following set of rules, where x and y are vertices, and φ and ψ are FO formulas:

$$x = y \mid \text{adj}(x, y) \mid \varphi \vee \psi \mid \varphi \wedge \psi \mid \neg\varphi \mid \exists x\varphi \mid \forall x\varphi.$$

As an example, triangle-freeness can be formulated as

$$\varphi = \neg\exists x_1\exists x_2\exists x_3(\text{adj}(x_1, x_2) \wedge \text{adj}(x_2, x_3) \wedge \text{adj}(x_3, x_1)).$$

The formula above assumes simple graphs (i.e., no loops nor multiple edges). If the graphs may have loops, then one should add the predicate $\neg(x_i = x_j)$ to the formula for every $i \neq j$.

The *monadic second-order* logic (MSO) extends FO by allowing quantification on *sets* of vertices and edges, with the incidence predicate $\text{inc}(v, e)$ indicating whether vertex v is incident to edge e , and the membership (\in) predicate. For instance, acyclicity can be formulated as

$$\varphi = \neg\exists X \neq \emptyset \forall x \in X \exists y_1 \in X \exists y_2 \in X (\neg(y_1 = y_2) \wedge \text{adj}(x, y_1) \wedge \text{adj}(x, y_2)).$$

Note that $X \neq \emptyset$ can merely be written as $\exists x \in X$. Note also that acyclicity cannot be expressed in FO as the length of the potential cycle is unbounded, from which it follows that one cannot quantify on vertices only for expressing acyclicity, because one does not know *how many* vertices should be considered. On the other hand, since FO can express properties such as C_4 -freeness, which, as mentioned before, requires certificates on $\tilde{\Omega}(\sqrt{n})$ bits, there is no hope of establishing a *meta-theorem* about FO regarding compact certification in all graphs. Nevertheless, a breakthrough in the theory of distributed certification was recently obtained by Bousquet, Feuilloley, and Pierron [20], who showed that every MSO predicate admits a distributed certification scheme with $O(\log n)$ -bit certificates in the family of graphs with bounded *treedepth*.

Algorithmic Meta-Theorems

A vibrant line of research in sequential computing is the development of algorithmic meta-theorems. According to Grohe and Kreutzer [45], algorithmic meta-theorems assert that certain families of algorithmic problems, typically defined by some logical and combinatorial

conditions, can be solved efficiently under some suitable definition of this term. Such theorems play an essential role in the theory of algorithms as they reveal a profound interplay between algorithms, logic, and combinatorics. One of the most celebrated examples of a meta-theorem is Courcelle’s theorem, which asserts that graph properties definable in MSO are decidable in linear time on graphs of bounded treewidth [10]. For an introduction to this fascinating research area, we refer to the surveys by Kreutzer [52] and Grohe [44].

Bousquet, Feuilloley, and Pierron in [20] introduced the exploration of algorithmic meta-theorems in distributed computing. Their primary result in this direction is that any MSO formula can be locally *certified* on graphs with bounded treedepth using a logarithmic number of bits per node, which represents the golden standard in certification. This theorem has numerous consequences for certification – for more details, we refer to [20]. Notably, the FO property C_4 -freeness, and the MSO property non-3-colorability, which both necessitate certificates of polynomial size in general, can be certified with just $O(\log n)$ -bit certificates in graphs of bounded treedepth. Bousquet et al.’s result has been extended to more comprehensive classes of graphs, including graphs excluding a small minor [6], as well as graphs of bounded *treewidth*, and graphs of bounded *cliquewidth*. However, this extension comes at the cost of slightly larger certificates, of $O(\log^2 n)$ bits, as seen in [34] and [33], respectively.

With significant advances in developing meta-theorems for distributed *certification*, there’s a notable absence of similar results for distributed *decision*. It prompts a natural question: could such results be obtained for the round-complexity of CONGEST? More concretely, the fundamental inquiry that remains unaddressed by Bousquet et al.’s paper, and by the subsequent works regarding distributed certification of MSO predicates is:

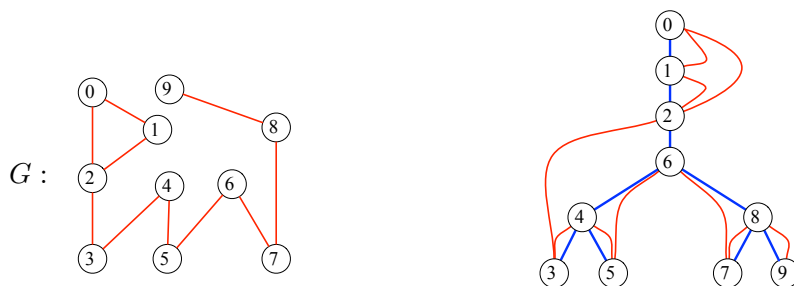
Question. *What is the round-complexity in CONGEST of deciding MSO formulas in graphs of bounded treedepth?*

A first step in answering this question was proposed in [59] where it is stated that, in any graph class of treedepth at most d , for every fixed connected graph H , H -freeness can be decided in $O(1)$ rounds in CONGEST. In this paper, we offer a comprehensive answer to the question. To elucidate our results, we first need to define the treedepth of a graph.

Treedepth

For any non-negative integer d , a (connected) graph G has treedepth at most d if there exists a rooted tree T spanning the vertices of G , with depth at most d , such that, for every edge $\{u, v\}$ in G , u is an ancestor of v in T , or v is an ancestor of u in T , see Fig. 1. The treedepth of a graph G , denoted by $\text{td}(G)$, is the smallest d for which such a tree exists.

The class of graphs with bounded treedepth, i.e., of treedepth d for some fixed $d \geq 0$, has strong connections with minor-closed families of graphs. Specifically, for any family \mathcal{F} of graphs closed under taking graph minors, the graphs in \mathcal{F} have bounded treedepth if and only if \mathcal{F} does not include all the paths [58]. Similarly, the graphs with bounded treedepth have a finite set of forbidden induced subgraphs, and any property of graphs monotonic with respect to induced subgraphs can be tested in polynomial time on graphs of bounded treedepth [58]. Computing the treedepth of a graph is NP-hard, but since treedepth is monotonic under graph minors, it is fixed-parameter tractable (FPT) [36]. Last but not least, MSO and FO have the same expressive power in graph classes of bounded treedepth [14].



■ **Figure 1** Embedding of a graph G into a tree T of depth 6.

1.1 Our Results

We prove that, for every MSO formula φ , there is an algorithm \mathcal{A} that, for every n -node graph G , decides whether $G \models \varphi$ in $O(2^{2^{\text{td}(G)}})$ rounds in the CONGEST model. That is, the round-complexity of \mathcal{A} depends only on the treedepth of the input graph, and on the MSO formula, i.e., it does not depend on the size n of the graph. Thus \mathcal{A} performs a constant number of rounds in any class of graphs with bounded treedepth. In particular, deciding non-3-colorability can be done in $O(1)$ rounds in graphs of bounded treedepth, in contrast to general graphs, for which deciding non-3-colorability requires a polynomial number of rounds by [43].

Our meta-theorem is essentially the best that one may hope regarding distributed model checking MSO formulas in a constant number of rounds in CONGEST. Indeed, the FO predicate “there is at least one vertex of degree > 2 ” requires $\Omega(n)$ rounds to be checked in this class. Hence our theorem cannot be extended to graphs of bounded treewidth or bounded cliquewidth, actually not even to bounded pathwidth, and not even to the class $\mathcal{P} \cup \mathcal{B}$ where \mathcal{P} is the set of all paths, and \mathcal{B} is the set of all graphs composed of a path to which is attached a claw at one of its endpoints.

We also consider distributed model checking of *labeled* graphs. For instance, one can check whether a given set of vertices is a feedback vertex set, i.e., whether the graph obtained by removing this set of vertices is acyclic. For such a predicate, it is sufficient to add a unary predicate to the logical structure used to mark the nodes, say $\text{mark}(x) = \text{true}$ means that vertex x is in the set. Using this unary predicate, φ can express the fact that there are no cycles in G passing only through nodes x for which $\text{mark}(x) = \text{false}$. As another example, the fact that a graph is properly 2-colored can be expressed using two unary boolean predicates red and blue, as

$$\varphi = \left(\forall x (\text{red}(x) \vee \text{blue}(x)) \right) \wedge \left(\forall x, y \neg \left(\text{adj}(x, y) \wedge ((\text{red}(x) \wedge \text{red}(y)) \vee (\text{blue}(x) \wedge \text{blue}(y))) \right) \right).$$

Since we also deal with MSO, we can also label edges. For instance, one can check whether a given set of edges forms a spanning tree. Indeed, it is sufficient to introduce a unary predicate used to mark the edges: $\text{mark}(e) = \text{true}$ means that edge e is in the set. As for feedback vertex set, using this unary predicate, φ can express the fact that the set of marked edges is a spanning tree (i.e., every node is incident to at least one marked edge, and any two vertices are connected by a path of marked edges). We show that deciding MSO formulas on *labeled* graphs of bounded treedepth can be done in $O(1)$ rounds in the CONGEST model.

More generally, we also consider the *optimization* variants of decision problems expressible in MSO on graphs of bounded treedepth. For instance, an independent set can be expressed as an MSO formula with a free variable S , such as $\varphi(S) = \forall x \in S \forall y \in S \neg \text{adj}(x, y)$. Then,

$max\varphi$, i.e., maximum independent set, consists in, given any graph $G = (V, E)$, finding the largest set $S \subseteq V$ such that $G \models \varphi(S)$. We show that, for every MSO formula $\varphi(S)$ with free variable $S \subseteq V$ or $S \subseteq E$, there is an algorithm for graphs of bounded treedepth solving $max\varphi$ (and $min\varphi$) in a constant number of rounds in the CONGEST model. This constant is of the form $O(g(\text{td}(G), \varphi))$ for some function g . Thanks to the expressive power of MSO, our results yield algorithms with a constant number of rounds in the CONGEST model on graphs of bounded treedepth for numerous popular optimization problems including minimum vertex cover, minimum feedback vertex set, minimum dominating set, maximum independent set, maximum induced forest, maximum clique, maximum matching, minimum spanning tree, Hamiltonian cycle, cubic subgraph, planar subgraph, Eulerian subgraph, Steiner tree, disjoint paths, min-cut, minor and topological minor containment, rural postman, k -colorability, edge k -colorability, partition into k cliques, and covering by k cliques. We also extend our results to *counting* problems, such as counting triangles or perfect matchings. Due to space restriction we only present in details in this article the distributed model checking result, for unlabeled graphs. For the optimization and counting variants, and the general case of labeled graphs, we refer to the full paper [28].

Finally, we briefly discuss some applications of our results to much larger classes of graphs, namely graphs of bounded *expansion* (see [58] for an extended introduction). Graphs of bounded expansion include planar graphs, and more generally, all classes of graphs defined from excluding minor. It was shown [59] that, for every class \mathcal{G} of graphs with bounded expansion, and every positive integer p , there is an algorithm performing in $O(\log n)$ rounds under the CONGEST model that partitions the vertex set V of any graph $G = (V, E) \in \mathcal{G}$ into $f(p)$ parts $V_1, \dots, V_{f(p)}$ such that every collection V_{i_1}, \dots, V_{i_q} of at most p parts, $1 \leq q \leq p$, $\{i_1, \dots, i_q\} \subseteq \{1, \dots, f(p)\}$, induces a (not necessarily connected) subgraph of G with treedepth at most p . The function f solely depends on the considered class \mathcal{G} of bounded expansion. The vertex partitioning $V_1, \dots, V_{f(p)}$ is called a low treedepth decomposition with parameter p . Plugging in our techniques into this framework, we show that, for every connected graph H , H -freeness can be decided in $O(\log n)$ rounds under the CONGEST model in any class of graphs with bounded expansion. This result was claimed in [59] with no proofs. We provide that claim with a complete formal proof.

1.2 Other Related Work

The quest for efficient (sublinear) algorithms for solving classical graph problems in the CONGEST model dates back to the seminal paper by Garay, Kutten and Peleg [38], where an algorithm for constructing an MST was designed. Since then, a long series of problems have been addressed, such as connectivity decomposition [7], tree embeddings [41] k -dominating set [53], stienner trees [54], min-cut [40, 56], max-flow [39], shortest path [47, 55], among others. Additionally, algorithms tailored to specific classes of networks have also been developed: DFS for planar graphs [42], MST for bounded genus graphs [46], MIS for networks excluding a fixed minor [9], etc.

Distributed certification is a very vivid topic, and many results have appeared since the survey [21]. A handful of recent papers considered *approximate* variants of the problem, a la property testing [15, 16, 22]. In particular, it was shown that every monotone (i.e., closed under taking subgraphs) and summable (i.e., stable by disjoint union) property Π has a compact approximate certification scheme in any proper minor-closed graph class [18]. Other recent contributions dealt with augmenting the power of the verifier in certification schemes, which includes tradeoffs between the size of the certificates and the number of rounds of the verification protocol [24], randomized verifiers [35], quantum verifiers [29],

and several hierarchies of certification mechanisms, including games between a prover and a disprover [1, 23], interactive protocols [11, 49, 57], and even zero-knowledge distributed certification [2], and distributed quantum interactive protocols [37].

2 Treedepth and treewidth

Throughout the paper, trees (or forests) are considered as rooted. The *depth* of a tree is the number of vertices of a longest path from the root to a leaf. The depth of a forest is the maximum depth among its trees. Let us recall the definition of the treedepth. The interested reader can refer to the book of Nešetřil and Ossona de Mendez [58] for further insights.

► **Definition 1** (treedepth). *The treedepth of a graph $G = (V, E)$ is the minimum depth of a forest $T = (V, F)$ on the same vertex set as G , such that, for any edge $\{u, v\}$ of G , one of the endpoints is an ancestor of the other in the forest T . Such a forest T is also called elimination forest of G .*

Observe that if G is connected then the forest T in the definition above is actually a tree. Also, Definition 1 does not require the forest T to be a subgraph of G . The following statement is an alternative, equivalent definition for treedepth. This recursive definition implicitly provides a recursive construction of an elimination tree.

► **Lemma 2** ([58]). *The treedepth of a graph G is:*

$$\text{td}(G) = \begin{cases} 1 & \text{if } G \text{ has a unique vertex,} \\ 1 + \min_{v \in V(G)} \text{td}(G - v) & \text{if } G \text{ is connected,} \\ \max\{\text{td}(C) \mid C \text{ is a connected component of } G\} & \text{otherwise.} \end{cases}$$

On the other hand, tree decompositions and treewidth of graphs were introduced by Robertson and Seymour [61].

► **Definition 3** (treewidth). *A tree decomposition of a graph $G = (V, E)$ is a pair (T, B) where $T = (I, F)$ is a tree, and $B = \{B_i, i \in I\}$ is a collection of subsets of vertices of G , called bags, such that the following conditions hold:*

- For every vertex of G , there exists some bag containing this vertex;
- For every edge e of G there is some bag containing both endpoints of e ;
- For every $v \in V$, the set $\{i \in I : v \in B_i\}$ of bags containing v forms a connected subgraph of T .

The width of a tree decomposition is the maximum size of a bag, minus one. The treewidth of a graph G , denoted by $\text{tw}(G)$, is the smallest width of a tree decomposition of G .

It is known [58] that the treedepth of a graph is at least its treewidth. Given an elimination tree T of a graph G , we can define a canonical tree decomposition associated to this same tree, such that the width of the decomposition corresponds to the depth of T , minus one. The following lemma is a straightforward consequence of the definitions of elimination trees and of tree decompositions.

► **Lemma 4** (canonical tree decomposition). *Let $T = (V, F)$ be an elimination tree of depth d of a graph $G = (V, E)$. Let us associate to each node u of T a bag $B(u)$ containing u and all the ancestors of u in T . Then $T = (V, F)$, and the corresponding set of bags $(B_u)_{u \in V}$, form a tree-decomposition of G , of width $d - 1$.*

For instance, the treedepth of an n -vertex path P_n is exactly $\lceil \log(n+1) \rceil$ (see, e.g., [58]). The treedepth of a graph does not increase when we delete some of its edges or vertices. Thus graphs of treedepth d have no paths on 2^d vertices. This observation yields the following lemma.

► **Lemma 5.** *Let $T = (V, F)$ be an elimination tree of a graph $G = (V, E)$ with $F \subseteq E$. Then the depth of T is at most $2^{\text{td}(G)}$.*

Proof. Let $d = \text{td}(G)$. Assume, for the purpose of contradiction, that T has depth larger than 2^d . It follows that the longest path P in T from its root to a leaf contains at least 2^d vertices. The path P is also a path in G , so the treedepth of P is at most the treedepth of G , i.e., at most d . This is a contradiction with the fact that, for n -node paths, $\text{td}(P_n) = \lceil \log(n+1) \rceil$. ◀

3 Tree decompositions and w -terminal recursive graphs

Courcelle’s theorem [10] states that any property expressible in MSO can be decided in linear (sequential) time on graphs of bounded treewidth. We use an alternative proof of Courcelle’s theorem, by Borie, Parker, and Tovey [5]. Indeed, this proof provides us with an explicit dynamic programming strategy, which will be used in our distributed protocol.

Graphs of bounded treewidth can also be defined recursively, based on a graph grammar. Let w be a positive integer. A w -terminal graph is a triple (V, W, E) where $G = (V, E)$ is a graph, and $W \subseteq V$ is a *totally ordered* set of at most w distinguished vertices. Vertices of W are called the *terminals* of the graph, and we denote by $\tau(G)$ the number of its terminals. As the terminal set W is totally ordered, we can refer to the r th terminal, for $1 \leq r \leq w$. Moreover, since vertices are given distinct identifiers in CONGEST, one can view W as ordered by these identifiers.

The class of w -terminal *recursive* graphs is defined, starting from w -terminal *base* graphs, by a sequence of *compositions*, or *gluings*. A w -terminal base graph is a w -terminal graph of the form (V, W, E) with $W = V$. A composition f acts on two¹ w -terminal graphs, and produces a new w -terminal graph, as follows (see Figure 2 for an example², for $w = 2$).

The graph $G = f(G_1, G_2)$ is obtained by, first, making disjoint copies of the two graphs G_1 and G_2 , and, second, “gluing” together some terminals of G_1 and G_2 . In the gluing operation, each terminal of G_1 is identified with at most one terminal of G_2 . Formally, the gluing performed by f is represented by a matrix $m(f)$ having $\tau(G) \leq w$ rows, and two columns, with integer entries in $\{0, \dots, \tau(G)\}$. At row r of the matrix, $m_{r,c}(f)$ indicates which terminal of each graph G_c , $c \in \{1, 2\}$ is identified to the r th terminal of graph G . If $m_{r,c}(f) = 0$, then no terminal of G_c is identified to terminal r of G (in particular, if $m_{r,1}(f) = m_{r,2}(f) = 0$, then the r th terminal of G is a new vertex; Nevertheless, this situation will not occur in our construction). Every terminal of G_c is identified to at most one terminal of G , i.e., each non-zero value in $\{1, \dots, \tau(G_c)\}$ appears at most once in the column c of $m(f)$.

A simple but crucial observation is that the number of possible different matrices, and hence of different composition operations f , is bounded by a function of w . Indeed the size of each matrix is at most $2w$, and each entry of the matrix is an integer between 0 and w .

¹ The definition of [5] considers composition operations of arbitrary arity, i.e., they consider gluing on three or more graphs simultaneously, and they also consider a special gluing, on a single graph which allows to “forget” some terminals. Technically, all these operations can be replaced by operations of arity 2, and we only use arity 2 for the sake of simplifying the presentation.

² The figure is borrowed from [33] with the agreement of the authors.

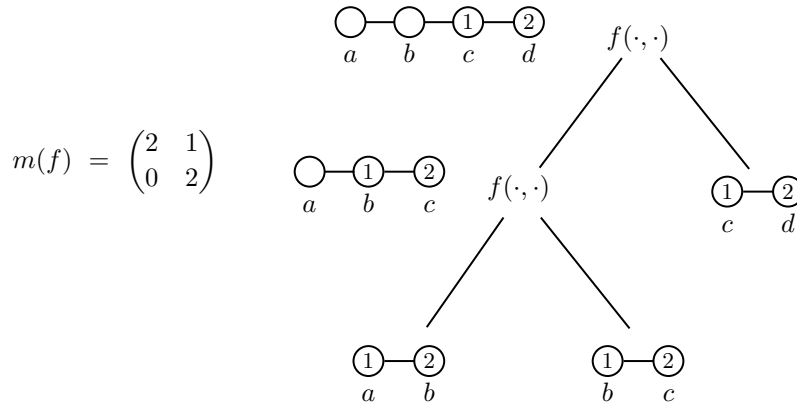


Figure 2 Paths as 2-terminal recursive graphs.

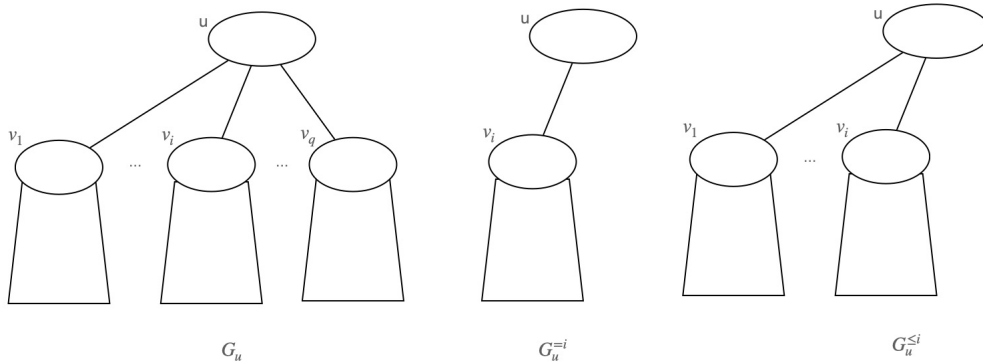


Figure 3 Tree-decompositions: graphs G_u , $G_u^{=i}$ and $G_u^{\leq i}$.

The class of w -terminal recursive graphs is exactly the class of graphs of treewidth at most $w - 1$ (see, e.g., Theorem 40 in [4]).

Let us briefly describe how a tree-decomposition of width $w - 1$ of a graph G can be transformed into a description of G as a w -terminal recursive graph. This construction will be crucial for efficiently deciding MSO properties of graph G . Let $T = (I, F)$ be a tree-decomposition of $G = (V, E)$ with bags of size at most w . The terminals correspond to the root bag. For every node u of T , we use the following notations, depicted in Figure 3:

- T_u is the subtree of T rooted at u ;
- B_u is the bag of node u , and $G_u^{\text{base}} = (G[B_u], B_u)$ is the w -terminal recursive base graph induced by bag B_u ;
- V_u is the union of all bags of T_u , and $G_u = (G[V_u], B_u)$ is the w -terminal graph induced by V_u , with B_u as set of terminals.

Let us now show that G_u is indeed a w -terminal recursive graph. This is clear when u is a leaf, since, in this case, $G_u = G_u^{\text{base}}$ is a base graph. Assume that u is not a leaf, and let v_1, \dots, v_q be the children of node u in T . The ordering of the children is arbitrary, but fixed. Let us introduce two new families of w -terminal recursive graphs as follows. Both are having B_u as set of terminals, and, for every $i \in \{1 \dots, q\}$:

- $G_u^{=i} = G[B_u \cup V_{v_i}]$, and
- $G_u^{\leq i} = G[B_u \cup V_{v_1} \cup \dots \cup V_{v_i}]$.

Observe that $G_u^{=i}$ is obtained by gluing G_{v_i} with the base graph G_u^{base} . More precisely,

$$G_u^{=i} = f_{(B_{v_i}, B_u)}(G_{v_i}, G_u^{base}), \quad (1)$$

where the gluing operation $f_{(B_{v_i}, B_u)}$ glues the terminals of $B_{v_i} \cap B_u$ of G_{v_i} to the corresponding terminals of B_u , and the new set of terminals is B_u . Also, for all $i \in \{1, \dots, q-1\}$, $G_u^{<i+1}$ is obtained by gluing $G_u^{<i}$ and $G_u^{=i+1}$ using the gluing function $f_{(B_u, B_u)}$ (which is the identity function on $|B_u|$ terminals), that is:

$$G_u^{<i+1} = f_{(B_u, B_u)}(G_u^{<i}, G_u^{=i+1}). \quad (2)$$

By construction, we get $G_u = G_u^{<q}$.

4 MSO logic and Courcelle's theorem

Recall that, using monadic second-order (MSO) logic formulas on graphs, we can express graph properties such as “G is not 3-colorable” or “G contains no triangles”. In order to solve optimization problems, we also consider MSO formulas with a free variable. That is, we consider formulas of the form $\varphi(S)$ where S is a set of vertices, or a set of edges. The corresponding optimization problem aims at finding a set S with maximum (or minimum, depending on the context) size satisfying $G \models \varphi(S)$. We may even assume that the vertices, or edges of the input graph $G = (V, E)$ have *polynomial* weights, that is the weight assignment $w : V \cup E \rightarrow \mathbb{Z}$ satisfies that, for every $x \in V \cup E$, $w(x)$ can be encoded with $O(\log n)$ bits. The problem $max\varphi$ then consists in computing the set S with maximum weight satisfying $G \models \varphi(S)$. In this framework we can express problems like maximum (weighted) independent set, minimum (weighted) dominating set, or minimum-weight spanning tree (MST).

4.1 Regular Predicates, Homomorphism Classes, and Composition

To start, let us first consider closed formulae only, i.e., with no free variable, and formulas with just one free (edge or vertex) set variable. (Further extensions are discussed in Section 6, with applications to labeled graphs and counting.) Using closed formulae, we can refer to graph predicates $\mathcal{P}(G)$, and, using formulas with free variables, we can refer to graph predicates $\mathcal{P}(G, X)$, where X denotes a subset of vertices or a subset of edges of G . For each possible assignment of X with corresponding values, \mathcal{P} is either true or false.

Any composition operation f over two w -terminal recursive graphs $G_1 = (V_1, W_1, E_1)$, and $G_2 = (V_2, W_2, E_2)$ naturally extends to a composition over pairs (G_1, X_1) , and (G_2, X_2) . If $G = f(G_1, G_2)$, we denote by \circ_f the composition over pairs. More precisely,

$$\circ_f((G_1, X_1), (G_2, X_2)) = (G, X),$$

the operation being valid only under some specific conditions. Let us consider the case when X_1 and X_2 are vertex-set variables. For each terminal t of G , if terminals from both G_1 and G_2 were mapped to t , say, terminals $t_1 \in W_1$ and $t_2 \in W_2$, then either $t_1 \in X_1$ and $t_2 \in X_2$, or $t_1 \notin X_1$ and $t_2 \notin X_2$. The set X is interpreted as the union of X_1 and X_2 , by identifying pairs of terminal vertices $t_1 \in X_1$ and $t_2 \in X_2$ that have been mapped on a same terminal of G . For edge-sets, the set X can also be seen as the union of two sets X_1 and X_2 , up to gluing the vertices specified by f . We refer to [5] for a description of the gluing operation, and of the interpretation of the values of the variables.

► **Definition 6** (regular predicate). *A graph predicate $\mathcal{P}(G, X)$ is regular if, for any value w , we can associate to w*

■ *a finite set \mathcal{C} of homomorphism classes,*

- an homomorphism function h , assigning to each w -terminal recursive graph G , and to any subset X of vertices or edges of G , a class $h(G, X) \in \mathcal{C}$, and
 - an update function $\odot_f : \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$ for each composition operation f ,
- such that:

1. If $h(G_1, X_1) = h(G_2, X_2)$ then $\mathcal{P}(G_1, X_1) = \mathcal{P}(G_2, X_2)$;
2. For any two w -terminal recursive graphs G_1 and G_2 , and any two sets X_1 and X_2 ,

$$h\left(\odot_f((G_1, X_1), (G_2, X_2))\right) = \odot_f(h(G_1, X_1), h(G_2, X_2)).$$

A class $c \in \mathcal{C}$ is said to be an *accepting class* if there exists (G, X) such that $h(G, X) = c$, and $\mathcal{P}(G, X)$ is true. By definition, the predicate \mathcal{P} holds for every (G', X') such that $h(G', X') = c$. A non accepting class c is called a *rejecting class*. The same definitions applies to predicates $\mathcal{P}(G)$, with no free variables.

Remark. Without loss of generality, we may assume that, in Definition 6, the class $c = h(G, X)$ with $G = (V, W, E)$ encodes the intersection of X with $G[W]$. Indeed, since W is of constant size, if X is a vertex-set, then we can add the set of all the ranks of the elements in $X_j \cap W$, with respect to the totally ordered set W , to the class c . And if X_j is an edge-set, then we can store each edge of X_j contained in $G[W]$ as the pair of ranks of its endpoints. In particular, we can assume that we are given a function $\text{Selected}(c, W)$ which, given a class c , and a set of terminals W , returns the unique intersection of X with the vertices, or the edges, of $G[W]$.

► **Theorem 7** ([5]). *Any predicate $\mathcal{P}(G, X)$ expressible by an MSO formula $\varphi(X)$ is regular. Moreover, given the formula $\varphi(X)$ and a parameter w , one can compute the set of classes \mathcal{C} , the update functions \odot_f over all possible composition operations f , as well as the homomorphism classes $h(G, X)$ for all base graphs G , and all possible values of variable X . (The same holds for predicates $\mathcal{P}(G)$ corresponding to closed formulas φ .)*

Let us emphasize that the width parameter w , and the formula φ in Theorem 7 are *constants*. Therefore, the set of homomorphism classes \mathcal{C} is of constant size, and can be computed, as well as functions \odot_f and homomorphism classes of base graphs, in *constant* time. This constant just depends on w and on φ .

4.2 Sequential Model-Checking and optimization

We have all ingredients to describe the sequential model-checking algorithm on graphs of bounded treewidth, that we will later use for designing our distributed protocol. In a nutshell, the algorithm proceeds by dynamic programming, from the leaves of the decomposition-tree to the root. When considering a node u , the program deals with the graph G_u induced by all bags in the subtree rooted at u , which is viewed as a w -terminal recursive graph with labels B_u . The program computes the homomorphism class of $h(G_u)$ using merely the homomorphism classes of its children, the bags of its children, and the subgraph of G induced by the bag of u .

► **Lemma 8** (bottom-up decision). *Let $\mathcal{P}(G)$ be a regular predicate on graphs, corresponding to a formula φ with no free variables. Let G be a graph, and let $T = (I, F)$ be a tree-decomposition of G with bags $\{B_u \mid u \in I\}$. Let u be a node of the tree decomposition, with children v_1, \dots, v_q for $q \geq 0$. The homomorphism class of $h(G_u)$ can be computed using only G_u^{base} , the values of B_{v_i} and $h(G_{v_i})$ for all $i \in \{1, \dots, q\}$.*

Proof. Observe that $h(G_u^{\text{base}})$ can be computed directly by Theorem 7. In particular this settles the case when u is a leaf. If u is an internal node, then, for each $i = 1, \dots, q$, we can compute $h(G_u^{=i})$ using $B_{v_i}, B_u, h(G_{v_i})$ and $h(G_u^{\text{base}})$ as follows. By Equation 1, we have $G_u^{=i} = f_{(B_{v_i}, B_u)}(G_{v_i}, G_u^{\text{base}})$. Since B_{v_i} and B_u are known, one can construct the function $f_{(B_{v_i}, B_u)}$, and, thanks to Theorem 7, one can retrieve the function $\odot_{f_{(B_{v_i}, B_u)}}$. By Definition 6, $h(G_u^{=i}) = \odot_{f_{(B_{v_i}, B_u)}}(h(G_{v_i}), h(G_u^{\text{base}}))$. Since the parameters on the right-hand side of the equality are known, one can compute $h(G_u^{=i})$.

Let us now show how to compute the values $h(G_u^{\leq i})$. For $i = 1$, $G_u^{\leq 1} = G_u^{=1}$, and thus $h(G_u^{\leq 1}) = h(G_u^{=1})$. For every $i \geq 2$, one can compute $h(G_u^{\leq i})$ using $B_u, h(G_u^{\leq i-1})$, and $h(G_u^{=i})$. Indeed, by Equation 2, $G_u^{\leq i} = f_{(B_u, B_u)}(G_u^{\leq i-1}, G_u^{=i})$, so again we have $h(G_u^{\leq i}) = \odot_{f_{(B_u, B_u)}}(h(G_u^{\leq i-1}), h(G_u^{=i}))$, and all parameters on the right-hand side have been computed previously.

Eventually, since $G_u = G_u^{\leq q}$, we get $h(G_u) = h(G_u^{\leq q})$. ◀

Algorithm 1 simultaneously presents the model-checking of regular predicates on graphs, and the optimization protocol for predicates on graphs and sets.

For the decision problem, we simply compute bottom-up, for each node u , the class of $h(G_u)$ using Lemma 8. At the root r , the algorithm accepts if $h(G_r)$ is an accepting class.

■ **Algorithm 1** Sequential decision for regular property \mathcal{P} on graph G .

Require: tree decomposition $T = (V_E, F_T)$ of width $w - 1$ of G ; formula φ and the corresponding homomorphism classes \mathcal{C} , homomorphism function h on base graphs, composition functions \odot_f ▷ See Theorem 7

- 1: **Bottom-up phase** on tree T , computes $h(G_u)$ for each node u :
 - 2: **for** each node u of T from the bottom to the root **do**
 - 3: let v_1, \dots, v_q be the children of u in T ▷ These nodes have been treated before u
 - 4: Compute $h(G_u)$ from $G_u^{\text{base}}, B_{v_i}$ and $h(G_{v_i}), 1 \leq i \leq q$ using Lemma 8
 - 5: **end for**
 - 6: **Decision at the root r :**
 - 7: Return true if $h(G_r)$ is an accepting class, otherwise return false.
-

A similar result can be obtained for regular predicates $\mathcal{P}(G, X)$, over (polynomially weighted) graphs and vertex sets. In order to compute the maximum or minimum weight set X such that $\mathcal{P}(G, X)$ holds, we can compute by dynamic programming, at each node u of the decomposition tree, a table $\text{OPT}(G_u)$ with $|\mathcal{C}|$ entries, one for each homomorphism class $c \in \mathcal{C}$ of \mathcal{P} , where $\text{OPT}(G_u)[c]$ corresponds to the optimum size of a partial solution set X_u such that $\mathcal{P}(G_u, X_u) = c$. This solution X_u is obtained by glueing partial solutions over the children nodes of u , and function **Selected** is used to avoid overcounting vertices in the intersection of partial solutions; see the full version of the paper [28] for complete details.

5 Distributed construction of the elimination tree

Our CONGEST protocol constructing an elimination tree of depth smaller than 2^d for graphs of treedepth at most d is depicted in Algorithm 2. A similar approach was previously used in in [59] for so-called small treedepth decompositions (see also Section 7 of this article).

■ **Algorithm 2** CONGEST algorithm computing an elimination tree of G in $O(2^{2d})$ rounds.

Require: Protocol $\text{leader}(G = (V, E), U)$ with a set $U \subseteq V$ of distinguished vertices (i.e., each vertex knows whether it belongs to U). After $O(\text{diam}(G))$ CONGEST rounds, each vertex u will know $\text{leader}(u)$, the minimum identifier in the component of $G[U]$ containing u .

```

1: Apply leader on all vertices of  $G$ 
2: Let  $r$  be the unique node such that  $r = \text{leader}(r)$            ▷ Unique since  $G$  is connected
3: Set  $\text{parent}(r) = r$                                            ▷  $r$  is the root of the tree
4: Mark vertex  $r$                                                ▷ Marked vertices are those already placed in the tree
5: Set  $\text{depth}(r) = 1$ 
6: for step  $i = 2$  to  $D = 2^d - 1$  do
7:   // At step  $i$  we identify the nodes of  $T$  of depth  $i$ 
8:   Apply leader on all unmarked vertices of  $G$                  ▷  $O(2^d)$  rounds
9:   Each unmarked vertex  $u$  broadcasts  $(\text{leader}(u), u)$  to its neighbours   ▷ One round
10:  for each marked vertex  $v$  of depth  $i$  do                       ▷ All in a same round
11:    for each  $\ell$  among values  $\text{leader}(u)$  received by  $v$  do
12:       $v$  picks the corresponding  $u(\ell)$  of minimum id ▷  $u(\ell)$  is a new node of depth  $i$ 
13:       $v$  adds  $u(\ell)$  to the list of its children
14:       $v$  sends to  $u(\ell)$  a message with its id indicating that it becomes its parent.
15:    end for
16:  end for
17:  for each vertex  $u$  that receives such a message from some  $v$  do ▷ All in a same round
18:     $u$  sets  $\text{parent}(u) = v$ ,  $\text{depth}(u) = i$  and marks itself
19:  end for
20: end for
21: if some vertex  $u$  is still unmarked then
22:    $u$  rejects because  $\text{td}(G) > d$            ▷  $G$  contains a path with more than  $D$  vertices
23: end if

```

► **Lemma 9.** *Let $G = (V, E)$ be the (connected) input network, and let $d \geq 1$ be an integer. There exists an algorithm performing in $O(2^{2d})$ rounds in CONGEST that outputs either an elimination tree $T = (T, F)$ of G with depth at most 2^d , or reports that $\text{td}(G) > d$. In the former case, each node $u \in V$ knows its parent and its children in the tree T at the end of the algorithm, as well as the depth of T .*

Proof. Algorithm 2 constructs an elimination tree following the same approach as Lemma 2, in a greedy manner: in a nutshell, it computes a depth-first search tree of G , heavily relying on the fact that, if G is of treedepth at most d , then all its paths are of length at most 2^d . Since G is connected, it starts with a root vertex $v = r$ (chosen arbitrarily), and then constructs elimination trees of $G \setminus v$, by treating each component separately. The components of $G \setminus v$ are identified by their leader with the smallest node's identifier of the component. Each unmarked node eventually knows its leader (Instruction 8). For a component with leader ℓ , we choose as root of the component a vertex that is adjacent to v (Instruction 12). In particular, every edge of the tree is also an edge of G (see Instruction 14).

The construction preserves the following invariant. The tree constructed after step i is an elimination tree of the subgraph induced by the marked vertices. Moreover, for each connected component of unmarked vertices, its outgoing edges are solely incident to a path from the root and a vertex v of depth i . In particular, at the end, T is an elimination tree

of G . Furthermore, T is a subtree of G . Therefore, by Lemma 5, if $\text{td}(G) \leq d$ then the depths of T is smaller than 2^d as requested, and the algorithm marks all vertices in less than 2^d phases. Consequently, if some vertices remain unmarked after this many rounds (Instruction 21), we correctly assert that $\text{td}(G) > d$.

Regarding the round-complexity, observe that, at each step, there is a call to Algorithm `leader` on the set of unmarked nodes (see, e.g., [48] for a detailed description of a leader-election algorithm). Its round complexity is $O(\text{diam}(G))$. The diameter of G is $O(2^{\text{td}(G)})$, and thus is it at most $O(2^d)$ (we can adapt algorithm `leader` such that, if it does not succeed in $O(2^d)$ rounds, then it rejects, which is correct as, in this case, $\text{td}(G) > d$). ◀

► **Lemma 10.** *Let $G = (V, E)$ be the input network, and let us assume that an elimination tree $T = (F, V)$ of depth smaller than 2^d has been constructed as in Lemma 9. There is a CONGEST algorithm constructing the canonical tree decomposition $(T, (B_u)_{u \in V})$ in $O(2^d)$ rounds. At the end of the algorithm, each node u knows its bag B_u as well as the graph $G[B_u]$ induced by the bag.*

Proof. The algorithm proceeds top-down. For each round $i = 1, \dots, D = 2^d - 1$, every node u at depth i computes B_u and $G[B_u]$. Observe that when u is the root, B_u is a singleton so $G[B_u]$ is trivial. If u is not the root, then u has received B_v and $G[B_v]$ from its parent v . Observe that $B_u = B_v \cup \{u\}$ and the edges of $G[B_u]$ are the edges of $G[B_v]$, plus the edges incident to u . Therefore, node u is able to compute the information from its parent, and to transmit it to its children. ◀

6 Distributed model checking and optimization

We have now all ingredients to prove our main result.

► **Theorem 11** (Distributed decision and optimization).

- For any closed MSO formula φ , there exists an algorithm which, for any n -node graph G , and any $d \geq 0$, decides whether $G \models \varphi$, or reports “large treedepth” if $\text{td}(G) > d$, running in $O(2^{2^d})$ rounds in the CONGEST model.
- For any MSO formula $\varphi(S)$ with a free variable S representing a vertex-set, or an edge-set, there exists an algorithm which, for any n -node graph G , and any $d \geq 0$, selects a set S of maximum weight satisfying $G \models \varphi(S)$, or reports “large treedepth” if $\text{td}(G) > d$, running in $g(d, \varphi)$ rounds in the CONGEST model for some function g .

Proof. By Lemmas 9 and 10, one can construct a canonical tree decomposition $T = (V, F)$ of $G = (V, E)$, with bags $\{B_u \mid u \in V\}$ of width at most 2^d (or correctly reject because $\text{td}(G) > d$), in $O(2^{2^d})$ rounds. Moreover each node u knows its parent $\text{parent}(u)$, its bag B_u , the graph $G[B_u]$, and its depth in T . By construction, the tree T is a subgraph of G . It remains to show that, based on these elements, one can implement the sequential algorithm (cf. Algorithm 1) in CONGEST.

Let us first consider model-checking of a closed formula φ . We describe how the bottom-up phase, and the decision at the root in Algorithm 1 can be implemented in $\text{depth}(T)$ rounds. Let us consider all steps $j \in \{1, \dots, \text{depth}(T)\}$, where each step consists of a single round. At step j , all nodes u of depth $k = \text{depth}(T) - j + 1$ can compute the homomorphism classes $h(G_u)$ in parallel (k decreases from $\text{depth}(T)$ to 1), and can send the results of this computation to their parents. Indeed, if u of depth k is a leaf, then it has all information needed to compute $h(G_u)$ already, because it only needs to know graph $G_u^{\text{base}} = G[B_u]$ (see Instruction 4 of Algorithm 1, and Lemma 8). If u it is not a leaf, then it also needs the bags

B_{v_i} , and the homomorphism classes $h(G_{v_i})$ from all its children $v_i, \{1, \dots, q\}$. But, at step j , node u has precisely already received these information from its children, who have sent them at the previous step $j - 1$. The decision at the root can be performed at round 1. The root accepts or rejects depending on its homomorphism class, as in Algorithm 1, and all other nodes accept. Therefore, if $G \models S$, then all nodes accept, otherwise the root rejects. Note that each message consists of a homomorphism class, thus the size of the messages is a constant. More precisely, messages are of size $\log |\mathcal{C}|$ bits, where \mathcal{C} denotes the set of homomorphism classes for property φ and treedepth at most 2^d .

For the optimization version, due to space restriction the details are given in the full version [28]. In a nutshell, the bottom up phase computes at each node u a table $\text{OPT}(G_u)$ of size \mathcal{C} which, for each homomorphism class $c \in \mathcal{C}$, stores the size of the optimal partial solution S_u corresponding to that class. Therefore the round complexity depends on the size of this table OPT , which is upper bounded by some function g depending on d and formula φ . A top-down phase is needed in order to mark the vertices of the global optimal solution. ◀

7 Applications to H -freeness for graphs of bounded expansion

For the many alternative definitions of graphs of bounded expansion, we refer to the book of Nešetřil and Ossona de Mendez in [58]. In terms of applications, we simply recall that the class of planar graphs, and, more generally, every class of graphs excluding a fixed minor, are classes of graphs of bounded expansion. It is known that graphs of bounded expansion admit so-called *low treedepth decompositions*.

► **Theorem 12** ([58]). *Let \mathcal{G} be a class of graphs of bounded expansion. There is a function $f : \mathbb{N} \rightarrow \mathbb{N}$ such that, for every integer $p > 0$, and every graph $G = (V, E) \in \mathcal{G}$, the vertex set of G can be partitioned into at most $f(p)$ parts $V_1, \dots, V_{f(p)}$ such that the union of any q parts, $1 \leq q \leq p$, induces a subgraph of G with treedepth at most q .*

A partition satisfying the property of Theorem 12 is called a *low treedepth decomposition* of G for parameter p . Interestingly, low treedepth decompositions can be efficiently computed in CONGEST, i.e., each vertex can compute the index $i \in \{1, \dots, f(p)\}$ of the part to which it belongs.

► **Theorem 13** ([59]). *For every graph class \mathcal{G} with bounded expansion, and every positive integer p , a low treedepth decomposition of G for parameter p can be computed in $O(\log n)$ rounds in CONGEST.*

The constant hidden in the big-O notation in the statement of Theorem 13 depends on the class \mathcal{G} and on the parameter p , and it is quite huge. The proof of Theorem 13 is sophisticated, but the algorithm is actually quite simple. It is merely based on the fact that graphs with bounded expansion have bounded degeneracy, and on the use of standard distributed tools for approximating the degeneracy of a graph in CONGEST. Combining Theorem 11 with Theorem 13, we can establish the following.

► **Corollary 14.** *Let \mathcal{G} be a class of graphs with bounded expansion, and let H be a connected graph. Deciding H -freeness for graphs in \mathcal{G} can be achieved $O(\log n)$ rounds under the CONGEST model.*

Proof. The algorithm works as follows. Let p be the number of vertices of H . First, compute a low treedepth decomposition $V_1, \dots, V_{f(p)}$ of the input graph $G = (V, E)$ into $f(p)$ parts for parameter p using Theorem 13. Then, for every non-empty set $I \subseteq [f(p)]$ with $|I| \leq p$,

let $G_I = G[\cup_{i \in I} V_i]$ be the graph induced by the parts V_i , $i \in I$. Note that there are at most $\binom{f(p)}{p}$ such subsets I , that is, a constant number of choices for I . Also observe that if a copy of H exists in graph G , then this copy of H belongs to at least one of the graphs G_I . Indeed, this holds for every set I of at most p parts V_i such that each of the p vertices of the copy of H is contained in some part V_i . Informally, we “guess” the colors of the copies of H , and I must contain these at most p colors. It is therefore sufficient to run the algorithm in Theorem 11 on each graph G_I in parallel, and to reject if one of the parallel executions finds a copy of H . This is doable because (1) G_I is of treedepth at most p , (2) if a copy of H exists, then it will be found in a connected component of G_I , thanks to the fact that H is connected, and (3) the property “ G_I is H -free” can be expressed as an MSO formula (actually, even as an FO formula), with p variables, one for each vertex of H . For instance, for a graph $H = (V_H, E_H)$ with $V_H = \{1, 2, \dots, p\}$, we can use the formula

$$\varphi_H = \neg \exists x_1, x_2, \dots, x_p \left(\bigwedge_{\{i,j\} \in E_H} \text{adj}(x_i, x_j) \right) \wedge \left(\bigwedge_{\{i,j\} \notin E_H} \neg \text{adj}(x_i, x_j) \right).$$

Thus the algorithm rejects if and only if the input graph contains a copy of H , as desired. ◀

In particular, Corollary 14 proves that H -freeness can be solved in $O(\log n)$ rounds in planar graphs under CONGEST. In contrast, for arbitrary graphs, even C_4 -freeness requires $\Omega(\sqrt{n})$ rounds, and, for every $p \geq 2$, there are $O(p)$ -vertex graphs H for which H -freeness requires $\Omega(n^{2-1/p})$ rounds [27]. Note that H -freeness can be considered in the usual sense (i.e., the input graph does not contain any copy of H as an induced subgraph), but also in the mere sense that there are no copies of H as a (non necessarily induced) subgraphs, by a straightforward adaptation of the MSO formula describing the problem.

8 Conclusion

In this paper, we established a meta-theorem about MSO formulas on graphs with bounded treedepth within the CONGEST model. Treedepth plays a fundamental role in the theory of sparse graphs of Nešetřil and Ossona de Mendez [58]. In particular, decomposing a graph in graphs of bounded treedepth is the crucial step in deriving a linear-time model-checking algorithm for FO on graphs of bounded *expansion* in the sequential computational model. Graphs of bounded expansion contain bounded-degree graphs, planar graphs, graphs of bounded genus, graphs of bounded treewidth, graphs that exclude a fixed minor, etc. Model-checking for FO on graphs of bounded expansion cannot be achieved in the CONGEST model since, as we already mention, checking an FO predicate as simple as “there is at least one vertex of degree > 2 ” requires $\Omega(n)$ rounds in n -node trees. Nevertheless, there might exist some fragments of FO that could be tractable on graphs of bounded expansion in the distributed setting. It would be interesting to identify the exact boundaries of intractability in this context, regarding both distributed decision, and distributed certification. An initial step in this direction was taken by Nešetřil and Ossona de Mendez in [59], resulting in a distributed algorithm for computing a low treedepth decomposition of graphs of bounded expansion, running in $O(\log n)$ rounds under CONGEST. As we illustrated, this results allows to efficiently decide FO-expressible decision problems (such as H -freeness, for H connected) for classes of graphs with bounded expansion, in $O(\log n)$ rounds. We restate the open question of [59]: Given a *local* FO formula $\varphi(x)$, i.e., a formula where $\varphi(x)$ depends on a fixed-radius neighborhood of vertex x only, can we mark all vertices satisfying φ in $O(\log n)$ rounds?

References

- 1 Alkida Balliu, Gianlorenzo D'Angelo, Pierre Fraigniaud, and Dennis Olivetti. What can be verified locally? *J. Comput. Syst. Sci.*, 97:106–120, 2018. doi:10.1016/J.JCSS.2018.05.004.
- 2 Aviv Bick, Gillat Kol, and Rotem Oshman. Distributed zero-knowledge proofs over networks. In *33rd ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2426–2458, 2022. doi:10.1137/1.9781611977073.97.
- 3 Lélia Blin, Laurent Feuilloley, and Gabriel Le Boudier. Optimal space lower bound for deterministic self-stabilizing leader election algorithms. *Discret. Math. Theor. Comput. Sci.*, 25:1–17, 2023. doi:10.46298/DMTCS.9335.
- 4 Hans L. Bodlaender. A partial k -arboretum of graphs with bounded treewidth. *Theoretical Computer Science*, 209(1-2):1–45, 1998. doi:10.1016/S0304-3975(97)00228-4.
- 5 Richard B. Borie, R. Gary Parker, and Craig A. Tovey. Automatic generation of linear-time algorithms from predicate calculus descriptions of problems on recursively constructed graph families. *Algorithmica*, 7(5&6):555–581, 1992. doi:10.1007/BF01758777.
- 6 Nicolas Bousquet, Laurent Feuilloley, and Théo Pierron. Local certification of graph decompositions and applications to minor-free classes. In *25th International Conference on Principles of Distributed Systems (OPODIS)*, volume 217 of *LIPICs*, pages 22:1–22:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPICs.OPODIS.2021.22.
- 7 Keren Censor-Hillel, Mohsen Ghaffari, and Fabian Kuhn. Distributed connectivity decomposition. In *Proceedings of the 2014 ACM symposium on Principles of distributed computing*, pages 156–165, 2014. doi:10.1145/2611462.2611491.
- 8 Keren Censor-Hillel, Ami Paz, and Mor Perry. Approximate proof-labeling schemes. *Theor. Comput. Sci.*, 811:112–124, 2020. doi:10.1016/J.TCS.2018.08.020.
- 9 Yi-Jun Chang. Efficient distributed decomposition and routing algorithms in minor-free networks and their applications. In *Proceedings of the 2023 ACM Symposium on Principles of Distributed Computing*, pages 55–66, 2023. doi:10.1145/3583668.3594604.
- 10 Bruno Courcelle. The monadic second-order logic of graphs. I. Recognizable sets of finite graphs. *Inf. Comput.*, 85(1):12–75, 1990. doi:10.1016/0890-5401(90)90043-H.
- 11 Pierluigi Crescenzi, Pierre Fraigniaud, and Ami Paz. Trade-offs in distributed interactive proofs. In *33rd International Symposium on Distributed Computing (DISC)*, volume 146 of *LIPICs*, pages 13:1–13:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPICs.DISC.2019.13.
- 12 Atish Das Sarma, Stephan Holzer, Liah Kor, Amos Korman, Danupon Nanongkai, Gopal Pandurangan, David Peleg, and Roger Wattenhofer. Distributed verification and hardness of distributed approximation. *SIAM J. Comput.*, 41(5):1235–1265, 2012. doi:10.1137/11085178X.
- 13 Andrew Drucker, Fabian Kuhn, and Rotem Oshman. On the power of the congested clique model. In *33rd ACM Symposium on Principles of Distributed Computing (PODC)*, pages 367–376, 2014. doi:10.1145/2611462.2611493.
- 14 Michael Elberfeld, Martin Grohe, and Till Tantau. Where first-order and monadic second-order logic coincide. *ACM Trans. Comput. Log.*, 17(4):25, 2016. doi:10.1145/2946799.
- 15 Gábor Elek. Planarity can be verified by an approximate proof labeling scheme in constant-time. *J. Comb. Theory, Ser. A*, 191:105643, 2022. doi:10.1016/J.JCTA.2022.105643.
- 16 Yuval Emek, Yuval Gil, and Shay Kutten. Locally restricted proof labeling schemes. In *36th International Symposium on Distributed Computing (DISC)*, volume 246 of *LIPICs*, pages 20:1–20:22. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPICs.DISC.2022.20.
- 17 Louis Esperet and Benjamin Lévêque. Local certification of graphs on surfaces. *Theor. Comput. Sci.*, 909:68–75, 2022. doi:10.1016/J.TCS.2022.01.023.
- 18 Louis Esperet and Sergey Norin. Testability and local certification of monotone properties in minor-closed classes. In *49th International Colloquium on Automata, Languages, and*

- Programming (ICALP)*, volume 229 of *LIPICs*, pages 58:1–58:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPICs.ICALP.2022.58.
- 19 Laurent Feuilloley. Introduction to local certification. *Discret. Math. Theor. Comput. Sci.*, 23(3):1–23, 2021. doi:10.46298/DMTCS.6280.
 - 20 Laurent Feuilloley, Nicolas Bousquet, and Théo Pierron. What can be certified compactly? compact local certification of MSO properties in tree-like graphs. In *41st ACM Symposium on Principles of Distributed Computing (PODC)*, pages 131–140, 2022. doi:10.1145/3519270.3538416.
 - 21 Laurent Feuilloley and Pierre Fraigniaud. Survey of distributed decision. *Bull. EATCS*, 119, 2016. URL: <http://eatcs.org/beatcs/index.php/beatcs/article/view/411>.
 - 22 Laurent Feuilloley and Pierre Fraigniaud. Error-sensitive proof-labeling schemes. *J. Parallel Distributed Comput.*, 166:149–165, 2022. doi:10.1016/J.JPDC.2022.04.015.
 - 23 Laurent Feuilloley, Pierre Fraigniaud, and Juho Hirvonen. A hierarchy of local decision. *Theor. Comput. Sci.*, 856:51–67, 2021. doi:10.1016/J.TCS.2020.12.017.
 - 24 Laurent Feuilloley, Pierre Fraigniaud, Juho Hirvonen, Ami Paz, and Mor Perry. Redundancy in distributed proofs. *Distributed Comput.*, 34(2):113–132, 2021. doi:10.1007/S00446-020-00386-Z.
 - 25 Laurent Feuilloley, Pierre Fraigniaud, Pedro Montealegre, Ivan Rapaport, Éric Rémila, and Ioan Todinca. Compact distributed certification of planar graphs. *Algorithmica*, 83(7):2215–2244, 2021. doi:10.1007/S00453-021-00823-W.
 - 26 Laurent Feuilloley, Pierre Fraigniaud, Pedro Montealegre, Ivan Rapaport, Éric Rémila, and Ioan Todinca. Local certification of graphs with bounded genus. *Discret. Appl. Math.*, 325:9–36, 2023. doi:10.1016/J.DAM.2022.10.004.
 - 27 Orr Fischer, Tzlil Gonen, Fabian Kuhn, and Rotem Oshman. Possibilities and impossibilities for distributed subgraph detection. In Christian Scheideler and Jeremy T. Fineman, editors, *Proceedings of the 30th on Symposium on Parallelism in Algorithms and Architectures, SPAA 2018, Vienna, Austria, July 16-18, 2018*, pages 153–162. ACM, 2018. doi:10.1145/3210377.3210401.
 - 28 Fedor V. Fomin, Pierre Fraigniaud, Pedro Montealegre, Ivan Rapaport, and Ioan Todinca. Distributed model checking on graphs of bounded treedepth. *CoRR*, abs/2405.03321, 2024. doi:10.48550/arXiv.2405.03321.
 - 29 Pierre Fraigniaud, François Le Gall, Harumichi Nishimura, and Ami Paz. Distributed quantum proofs for replicated data. In *12th Innovations in Theoretical Computer Science Conference (ITCS)*, volume 185 of *LIPICs*, pages 28:1–28:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPICs.ITCS.2021.28.
 - 30 Pierre Fraigniaud, Mika Göös, Amos Korman, Merav Parter, and David Peleg. Randomized distributed decision. *Distributed Comput.*, 27(6):419–434, 2014. doi:10.1007/S00446-014-0211-X.
 - 31 Pierre Fraigniaud, Mika Göös, Amos Korman, and Jukka Suomela. What can be decided locally without identifiers? In *32nd ACM Symposium on Principles of Distributed Computing (PODC)*, pages 157–165, 2013. doi:10.1145/2484239.2484264.
 - 32 Pierre Fraigniaud, Amos Korman, and David Peleg. Towards a complexity theory for local distributed computing. *J. ACM*, 60(5):35:1–35:26, 2013. doi:10.1145/2499228.
 - 33 Pierre Fraigniaud, Frédéric Mazoit, Pedro Montealegre, Ivan Rapaport, and Ioan Todinca. Distributed certification for classes of dense graphs. In *37th International Symposium on Distributed Computing (DISC)*, volume 281 of *LIPICs*, pages 20:1–20:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPICs.DISC.2023.20.
 - 34 Pierre Fraigniaud, Pedro Montealegre, Ivan Rapaport, and Ioan Todinca. A meta-theorem for distributed certification. *Algorithmica*, 86(2):585–612, 2024. doi:10.1007/S00453-023-01185-1.
 - 35 Pierre Fraigniaud, Boaz Patt-Shamir, and Mor Perry. Randomized proof-labeling schemes. *Distributed Comput.*, 32(3):217–234, 2019. doi:10.1007/S00446-018-0340-8.

- 36 Jakub Gajarský and Petr Hlinený. Kernelizing MSO properties of trees of fixed height, and some consequences. *Log. Methods Comput. Sci.*, 11(1):1–26, 2015. doi:10.2168/LMCS-11(1:19)2015.
- 37 François Le Gall, Masayuki Miyamoto, and Harumichi Nishimura. Distributed quantum interactive proofs. In *40th International Symposium on Theoretical Aspects of Computer Science (STACS)*, volume 254 of *LIPICs*, pages 42:1–42:21. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPICs.STACS.2023.42.
- 38 JA Garay, S Kutten, and D Peleg. A sub-linear time distributed algorithm for minimum-weight spanning trees. In *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*, pages 659–668. IEEE, 1993.
- 39 Mohsen Ghaffari, Andreas Karrenbauer, Fabian Kuhn, Christoph Lenzen, and Boaz Patt-Shamir. Near-optimal distributed maximum flow. In *Proceedings of the 2015 ACM Symposium on Principles of Distributed Computing*, pages 81–90, 2015.
- 40 Mohsen Ghaffari and Fabian Kuhn. Distributed minimum cut approximation. In *International Symposium on Distributed Computing*, pages 1–15. Springer, 2013. doi:10.1007/978-3-642-41527-2_1.
- 41 Mohsen Ghaffari and Christoph Lenzen. Near-optimal distributed tree embedding. In *International Symposium on Distributed Computing*, pages 197–211. Springer, 2014. doi:10.1007/978-3-662-45174-8_14.
- 42 Mohsen Ghaffari and Merav Parter. Near-optimal distributed dfs in planar graphs. In *31st International Symposium on Distributed Computing (DISC 2017)*, pages 21:1–21:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPICs.DISC.2017.21.
- 43 Mika Göös and Jukka Suomela. Locally checkable proofs in distributed computing. *Theory Comput.*, 12(1):1–33, 2016. doi:10.4086/TOC.2016.V012A019.
- 44 Martin Grohe. Logic, graphs, and algorithms. In *Logic and Automata: History and Perspectives, in Honor of Wolfgang Thomas*, volume 2 of *Texts in Logic and Games*, pages 357–422. Amsterdam University Press, 2008. URL: <https://eccc.weizmann.ac.il/report/2007/091/>.
- 45 Martin Grohe and Stephan Kreutzer. Methods for algorithmic meta theorems. In *Model Theoretic Methods in Finite Combinatorics - AMS-ASL Joint Special Session*, volume 558, pages 181–206. AMS, 2009. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.395.8282&rep=rep1&type=pdf>.
- 46 Bernhard Haeupler, Taisuke Izumi, and Goran Zuzic. Low-congestion shortcuts without embedding. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*, pages 451–460, 2016. doi:10.1145/2933057.2933112.
- 47 Monika Henzinger, Sebastian Krinninger, and Danupon Nanongkai. A deterministic almost-tight distributed algorithm for approximating single-source shortest paths. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 489–498, 2016. doi:10.1145/2897518.2897638.
- 48 Juho Hirvonen and Jukka Suomela. *Distributed Algorithms 2020*. Aalto University, 2020. URL: <https://jukkasuomela.fi/da2020/>.
- 49 Gillat Kol, Rotem Oshman, and Raghuvansh R. Saxena. Interactive distributed proofs. In *37th ACM Symposium on Principles of Distributed Computing (PODC)*, pages 255–264. ACM, 2018. URL: <https://dl.acm.org/citation.cfm?id=3212771>.
- 50 Amos Korman, Shay Kutten, and Toshimitsu Masuzawa. Fast and compact self-stabilizing verification, computation, and fault detection of an MST. *Distributed Comput.*, 28(4):253–295, 2015. doi:10.1007/S00446-015-0242-Y.
- 51 Amos Korman, Shay Kutten, and David Peleg. Proof labeling schemes. *Distributed Comput.*, 22(4):215–233, 2010. doi:10.1007/S00446-010-0095-3.
- 52 Stephan Kreutzer. Algorithmic meta-theorems. In *Finite and Algorithmic Model Theory*, volume 379 of *London Mathematical Society Lecture Note Series*, pages 177–270. Cambridge University Press, 2011. URL: <http://www.cs.ox.ac.uk/people/stephan.kreutzer/Publications/amt-survey.pdf>.

- 53 Shay Kutten and David Peleg. Fast distributed construction of k -dominating sets and applications. In *Proceedings of the fourteenth annual ACM symposium on Principles of distributed computing*, pages 238–251, 1995.
- 54 Christoph Lenzen and Boaz Patt-Shamir. Improved distributed steiner forest construction. In *Proceedings of the 2014 ACM symposium on Principles of distributed computing*, pages 262–271, 2014. doi:10.1145/2611462.2611464.
- 55 Danupon Nanongkai. Distributed approximation algorithms for weighted shortest paths. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 565–573, 2014. doi:10.1145/2591796.2591850.
- 56 Danupon Nanongkai and Hsin-Hao Su. Almost-tight distributed minimum cut algorithms. In *International Symposium on Distributed Computing*, pages 439–453. Springer, 2014. doi:10.1007/978-3-662-45174-8_30.
- 57 Moni Naor, Merav Parter, and Eylon Yogev. The power of distributed verifiers in interactive proofs. In *31st ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1096–1115. SIAM, 2020. doi:10.1137/1.9781611975994.67.
- 58 Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity - Graphs, Structures, and Algorithms*, volume 28 of *Algorithms and combinatorics*. Springer, 2012. doi:10.1007/978-3-642-27875-4.
- 59 Jaroslav Nešetřil and Patrice Ossona de Mendez. A distributed low tree-depth decomposition algorithm for bounded expansion classes. *Distributed Comput.*, 29(1):39–49, 2016. doi:10.1007/s00446-015-0251-x.
- 60 David Peleg. *Distributed computing: a locality-sensitive approach*. SIAM, 2000.
- 61 Neil Robertson and Paul D. Seymour. Graph minors. III. Planar tree-width. *J. Comb. Theory, Ser. B*, 36(1):49–64, 1984. doi:10.1016/0095-8956(84)90013-3.