

First-Order Model Checking on Structurally Sparse Graph Classes

Jan Dreier
TU Wien

dreier@ac.tuwien.ac.at

Nikolas Mählmann
University of Bremen

maehlmann@uni-bremen.de

Sebastian Siebertz
University of Bremen

siebertz@uni-bremen.de

Abstract

A class of graphs is **structurally nowhere dense** if it can be constructed from a nowhere dense class by a first-order transduction. Structurally nowhere dense classes vastly generalize nowhere dense classes and constitute important examples of monadically stable classes. **We show that the first-order model checking problem is fixed-parameter tractable on every structurally nowhere dense class of graphs.**

Our result builds on a recently developed game-theoretic characterization of monadically stable graph classes. As a second key ingredient of independent interest, we provide a polynomial-time algorithm for approximating weak neighborhood covers (on general graphs). We combine the two tools into a recursive locality-based model checking algorithm. This algorithm is efficient on every monadically stable graph class admitting flip-closed sparse weak neighborhood covers, where flip-closure is a mild additional assumption. Thereby, establishing efficient first-order model checking on monadically stable classes is reduced to proving the existence of flip-closed sparse weak neighborhood covers on these classes – a purely combinatorial problem. We complete the picture by proving the existence of the desired covers for structurally nowhere dense classes: we show that every structurally nowhere dense class can be sparsified by contracting local sets of vertices, enabling us to lift the existence of covers from sparse classes.

Contents

1	Introduction	1
2	Technical Overview	5
3	Preliminaries	9
3.1	Graphs	9
3.2	Logic	10
3.3	Stability	11
3.4	Flipper Game	12
3.5	Weak Neighborhood Covers	14
4	Guarded Formulas and Local Types	14
4.1	Guarded Formulas	14
4.2	Games	15
4.3	Local Games	16
4.4	Local Games Determine Global Games	16
4.5	Local Games and Local Types	19
4.6	Games and Types with Guards	20
5	Model Checking	22
5.1	Setup	22
5.2	Computing Guarded Formulas	23
5.3	Reducing the Evaluation Radius	29
5.4	Main Result	32
5.5	A Note on the Computability of $f(\varphi)$	33
6	Approximating Weak Neighborhood Covers	34
6.1	Linear Programming	34
6.2	ILP Formulation	35
6.3	Fractional Weak Neighborhood Covers	36
6.4	Randomized Sampling	37
6.5	Minimum Membership Set Cover Reduction	38
7	Weak Neighborhood Covers in Structurally Sparse Classes	40
7.1	Local Contractions and Weak Neighborhood Covers	40
7.2	Background on Structurally Sparse Graphs	41
7.3	Quasi-Bush Decompositions of Structurally Nowhere Dense Classes	44
7.4	Sibling Contractions in Quasi-Bushes	46
7.5	Outline of the Contraction Procedure	48
7.6	Phase One: Eliminating Coat Hangers	52
7.7	Phase Two: Bounding In and Out Sets	54
7.8	Sparsity of the Contraction	59
7.9	Wrapping Up	62

1 Introduction

Logic provides a versatile and elegant formalism for describing algorithmic problems. For example, the k -colorability problem (for every fixed k) and Hamiltonian path problem can be formulated in monadic second-order logic (MSO), while the k -independent set problem, k -dominating set problem, and many more, can be formulated in first-order logic (FO). Therefore, the tractability of the model checking problem for a logic, that is, the problem of deciding for a given structure and formula whether the formula is true in the structure, implies tractability for a whole class of problems, namely for all problems definable in that logic. For this reason, model checking results for logics are often called *algorithmic meta-theorems*. Due to the rich structure theory that graph theory offers, many algorithmic meta-theorems are formulated for graphs, but most of them can easily be extended to general relational structures. In the following, we will hence restrict our discussion to (colored) graphs and graph classes. Probably the best-known algorithmic meta-theorem is the celebrated theorem of Courcelle, stating that every graph property definable in MSO can be decided in linear time on every class of graphs with bounded treewidth [Cou90]. Due to their generality and wide applicability, algorithmic meta-theorems have received significant attention in contemporary research. We refer to the surveys [Gro08, GK11, Kre11].

Every fixed first-order formula with quantifier rank q can be tested on n -vertex graphs in time $n^{\mathcal{O}(q)}$ by a straight-forward recursive algorithm. In general this running time cannot be expected to be improved, as, for example, the k -independent set problem cannot be solved in time $f(k) \cdot n^{o(k)}$ for any function f , assuming the exponential time hypothesis (ETH) [IP01, IPZ01], and the k -independent set problem is expressible by an FO formula with quantifier rank k . This raises the question of classifying graph classes on which we can evaluate the truth of an FO formula φ in time $f(\varphi) \cdot n^{\mathcal{O}(1)}$ for some function f . Phrased in terms of parameterized complexity theory, the question is on which graph classes the FO model checking problem is *fixed-parameter tractable*¹. While the **classification of tractable MSO model checking is essentially complete** [GHL⁺14, KT10], we are far from a full classification of tractable FO model checking.

The starting point for such a classification on sparse graph classes is **Seese's result** that every FO property of graphs can be decided in linear time on every class of graphs of **bounded degree** [See96]. His result was extended to more and more general classes of sparse graphs [FG01a, FG01b, DGK07, DKT10, Kre11, GKS17]. The last result of Grohe, Kreutzer and Siebertz [GKS17] shows that every FO definable property of graphs is decidable in nearly linear time on every nowhere dense class of graphs. **Nowhere dense classes of graphs are very general classes of sparse graphs and turn out to be a tractability barrier for FO model checking on monotone classes** (that is, classes that are closed under taking subgraphs): if a monotone class of graphs is not nowhere dense, then testing first-order properties for inputs from this class is as hard as for general graphs [DKT10, Kre11].

Hence, the classification of tractable FO model checking on monotone classes is complete, which led to a shift of attention to hereditary graph classes (that is, classes that are closed under taking induced subgraphs), where model checking beyond sparse classes may be possible. One important result for classes of dense graphs is the extension of Courcelle's theorem for MSO to graphs of **bounded cliquewidth** [CMR00], which also implies tractability for FO on classes of **locally bounded cliquewidth**, for example, on map graphs [EK17]. Positive results were also obtained for some dense graph classes definable by geometric means, for example, for restricted subclasses of **interval graphs** [GHK⁺13], and for restricted subclasses of circular-arc, circle, box, disk, and polygon-visibility graphs [HPR17]. Many of these results for dense graph classes are generalized by the **recent breakthrough showing that FO model checking is fixed-parameter tractable on every class**

¹When the function f is not computable one also speaks of *non-uniform fixed-parameter tractability*. We will comment on the uniformity of our result in Section 5.5.

with bounded twin-width [BKTW21] if a witnessing contraction sequence is given as additional input. In fact, it turns out that when considering classes of ordered graphs, then classes of bounded twin-width define the tractability border for efficient FO model checking [BGdM⁺22].

In an attempt to extend the sparsity-based methods also to dense graph classes, researchers considered *structurally sparse graph classes*, which are defined as first-order transductions of sparse graph classes, see [GKN⁺20, NdM16]. Intuitively, a first-order transduction creates a graph H from a graph G on (a subset of) the vertices of G by coloring the vertices of G , replacing the edge relation by a first-order definable edge relation, and finally restricting the vertex set to a definable subset. We call a binary relation R definable in a graph G if there is a formula $\eta(x, y)$ such that $(u, v) \in R \Leftrightarrow G \models \eta(u, v)$. Similarly, a set U is definable if there is a formula $\nu(x)$ such that $v \in U \Leftrightarrow G \models \nu(v)$. We will formally define transductions in Section 3. Simple examples of first-order transductions are graph complementations and fixed graph powers. Another elementary and yet particularly important example of a transduction is that of a *flip*. Given two sets $A, B \subseteq V(G)$ that are each marked with a color, a flip complements the adjacency between vertices of A and B , which is easily definable by an atomic formula. We say that a class \mathcal{D} of graphs is *structurally nowhere dense* if there exists a nowhere dense class \mathcal{C} such that \mathcal{D} is a first-order transduction of \mathcal{C} . Note that structurally nowhere dense classes are vastly more general than nowhere dense classes, and in particular include classes of dense graphs. On the other hand, *structurally nowhere dense classes are incomparable to classes with bounded twin-width*.

As nowhere dense classes stand on top of the hierarchy of tractable sparse classes, *structurally nowhere dense classes arguably form the most general structurally sparse classes for which one could hope to achieve efficient model checking*. Until now, there was no indication on how to solve even special cases like the INDEPENDENT SET problem, let alone the full FO model checking problem on these classes. In this work, we set a milestone in this line of research by proving the following theorem.

Theorem 1. *For every structurally nowhere dense class \mathcal{C} there exists a function $f: \mathbb{N} \rightarrow \mathbb{N}$ such that, given a graph $G \in \mathcal{C}$ and sentence φ , one can decide whether $G \models \varphi$ in time $f(|\varphi|) \cdot |V(G)|^{11}$.*

Until now, efficient model checking algorithms could be lifted from tractable base classes to transductions of classes with bounded degree [GHO⁺20] and transductions of classes with bounded local cliquewidth [BDG⁺22]. In the latter result, the degree of the polynomial running time is non-uniform and depends on the class under consideration. For classes with structurally bounded expansion (every class of bounded expansion is nowhere dense but the notion of nowhere denseness is strictly more general than that of bounded expansion), efficient model checking is possible when additionally a special coloring is given with the input [GKN⁺20]. We remark that also the model checking algorithm on classes of unordered graphs of bounded twin-width requires a so-called contraction sequence as additional input and until now is only a conditional result [BKTW21].

At this point we want to highlight that our algorithm does not depend on any additional colorings or decompositions as part of the input and has a polynomial running time whose degree is independent of the class \mathcal{C} , setting it apart from the existing algorithms for classes with bounded twin-width, structurally bounded expansion and transductions of classes with bounded local cliquewidth.

Our result builds on notions from classical model theory. As observed in [AA14], *a monotone class of graphs is stable if and only if it is nowhere dense*. The notion of stability provides one of the most important dividing lines between wild and tame theories in model theory [She90]. Another important notion is *NIP*, which includes even larger graph classes. Surprisingly, it turns out that *on monotone classes all three notions of dependence, stability and nowhere denseness are*

equivalent. Important subclasses of stable and NIP classes are *monadically stable* and *monadically NIP* classes. These remain stable/NIP when graphs from the class are expanded by an arbitrary number of unary predicates (colors). All concepts relevant for this work will be formally defined in Section 3.

All hereditary classes on which the model checking problem is known to be fixed-parameter tractable are monadically NIP, **which led to the conjecture that a hereditary class of graphs admits fixed-parameter tractable model checking if and only if it is monadically NIP**, see for example [war16, GPT22]. It turns out that a hereditary class of *ordered* graphs is NIP if and only if it has bounded twin-width [BGODM⁺22]. Furthermore, a hereditary graph class is stable/NIP if and only if it is monadically stable/NIP [BL22], which further highlights the importance of these notions on hereditary graph classes. Every structurally nowhere dense class is monadically stable [PZ78, AA14] and in fact **it has been conjectured that the notions of monadic stability and structural nowhere denseness coincide**, see for example [GPT22].

Let us give a brief sketch of our approach to prove Theorem 1, which will allow us to highlight further contributions. By Gaifman’s Locality Theorem [Gai82], the problem of deciding whether a first-order sentence φ is true in a graph can be reduced to first testing whether other *local* formulas are true in the graph and then solving a colored variant of the (distance- d) independent set problem. The independent set problem itself can then also be reduced to the evaluation of local formulas in a coloring of the graph. Local formulas can be evaluated in bounded-radius neighborhoods of the graph, where the radius depends only on φ . Hence, whenever the local neighborhoods in graphs from a class \mathcal{C} admit efficient model checking, then one immediately obtains an efficient model checking algorithm for \mathcal{C} . This technique was first employed in [FG01b] and is also the basis of the model checking algorithm of [GKS17] on nowhere dense graph classes. A key ingredient on these classes was a characterization of these classes in terms of a recursive decomposition of local neighborhoods using a game called *Splitter game* [GKS17]. As the Splitter game can only decompose sparse graphs, **we replace it in our algorithm by a recently developed characterization of monadically stable classes via the so-called *Flipper game*** [GMM⁺23]. We will give a more detailed technical overview below, where we highlight in particular the many significant differences with the approach of [GKS17].

The algorithmic idea of [GKS17] is to reduce the evaluation of a formula φ to the evaluation of formulas in local neighborhoods and use the Splitter game decomposition to guide a recursion into local neighborhoods. Doing this naively would lead to a branching degree of n and a running time of $\Omega(n^\ell)$, where the depth ℓ of the recursion grows with φ . The solution presented in [GKS17] was to work with a rank-preserving normal form of formulas and to cluster nearby neighborhoods and handle them together in one recursive call using so-called *sparse r -neighborhood covers*. An r -neighborhood cover with *degree* d and *spread* s of a graph G is a family \mathcal{X} of subsets of $V(G)$, called *clusters*, such that the r -neighborhood of every vertex is contained in some cluster, every cluster has radius at most s , and every vertex appears in at most d clusters. We say a class \mathcal{C} admits *sparse neighborhood covers* if for every r and $\varepsilon > 0$ there exist constants $c(r, \varepsilon)$ and $\sigma(r)$ such that every $G \in \mathcal{C}$ admits an r -neighborhood cover with spread at most $\sigma(r)$ and degree at most $c(r, \varepsilon) \cdot |G|^\varepsilon$. It was proved in [GKS17] that every nowhere dense graph class admits sparse neighborhood covers. By scaling ε appropriately the recursive data structure needed for model checking can be constructed in the desired fpt running time.

Neighborhood covers are an important tool with applications, for example, in the design of distributed algorithms. We refer to [Pel00] for extensive background on applications and constructions of sparse neighborhood covers. As mentioned above, sparse neighborhood covers exist for nowhere dense graph classes [GKS17], and in fact for monotone classes their existence characterizes nowhere

dense classes [GKR⁺18]. On the other hand, sparse neighborhood covers are known not to exist for general graph classes: for every r and $s \geq 3$ there exist infinitely many graphs G for which every r -neighborhood cover of spread at most s has degree $\Omega(|G|^{1/s})$ [TZ05]. For monadically stable classes we neither know whether sparse r -neighborhood covers exist, nor, if they should exist, how to efficiently compute them.

For our model checking algorithm we may relax the assumptions on neighborhood covers and consider *weak* neighborhood covers. A *weak r -neighborhood cover* with *degree* d and *spread* s of a graph G is a family \mathcal{X} of subsets of $V(G)$, again called *clusters*, such that the r -neighborhood of every vertex is contained in some cluster, every cluster is a subset of an s -neighborhood in G (but does not necessarily have radius at most s itself), and every vertex occurs in at most d clusters. We say that a class \mathcal{C} admits *sparse weak neighborhood covers* if for every r and $\varepsilon > 0$ there exist constants $c(r, \varepsilon)$ and $\sigma(r)$ such that every $G \in \mathcal{C}$ admits a weak r -neighborhood cover with spread at most $\sigma(r)$ and degree at most $c(r, \varepsilon) \cdot |G|^\varepsilon$. It is easy to modify the proof of [TZ05] to show that general graph classes do not admit sparse weak neighborhood covers. However, note that the classes exhibited in [TZ05] are not monadically NIP. While we are still not able to prove that sparse weak neighborhood covers exist for monadically stable classes, as a main contribution of independent interest we prove that weak r -neighborhood covers can be efficiently approximated in general. We prove the following theorem.

Theorem 10. *There is an algorithm that gets as input an n -vertex graph G and numbers $r, s \in \mathbb{N}$ and computes in time $\mathcal{O}(n^{9.8})$ a weak r -neighborhood cover with degree $\mathcal{O}(\log(n)^2 + 1)d^*$ and spread s , where d^* is the smallest number such that G admits a weak r -neighborhood cover with degree d^* and spread s .*

Note that the logarithmic factors are negligible when we aim for neighborhood covers of degree $\mathcal{O}(n^\varepsilon)$. As a consequence of the theorem, when a class of graphs admits sparse weak neighborhood covers, then they are efficiently computable. With Theorem 10 established, we provide a clear path towards generalizing tractable model checking to all monadically stable graph classes, as made precise in the following theorem.

Theorem 5. *Let \mathcal{C} be a monadically stable graph class admitting flip-closed sparse weak neighborhood covers. There exists a function $f: \mathbb{N} \rightarrow \mathbb{N}$ such that, given a graph $G \in \mathcal{C}$ and sentence φ , one can decide whether $G \models \varphi$ in time $f(|\varphi|) \cdot |V(G)|^{11}$.*

Here, flip-closure is a mild closure condition requiring that each class \mathcal{C}_ℓ obtained by applying ℓ flips to graphs from \mathcal{C} admits sparse weak neighborhood covers with the same bound on the spread as \mathcal{C} . We expect that this closure property will be fulfilled naturally, as we are mainly interested in properties of graph classes that are closed under transductions, and thus also under any constant number of flips. With Theorem 5, the question of efficient model checking on monadically stable graph classes reduces to a purely combinatorial question about the existence of sparse weak neighborhood covers. We conjecture that the required neighborhood covers exist even for monadically NIP classes.

Conjecture 1. *Every monadically NIP class admits flip-closed sparse weak neighborhood covers.*

We conclude Theorem 1 from Theorem 5 and the following theorem, establishing that structurally nowhere dense classes admit flip-closed sparse weak neighborhood covers.

Theorem 11. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon)$ such that for every $G \in \mathcal{C}$ there exists a weak r -neighborhood cover with degree at most $c(r, \varepsilon) \cdot |G|^\varepsilon$ and spread at most $34r$. In particular, \mathcal{C} admits flip-closed sparse weak neighborhood covers.*

We believe that the methods to prove Theorem 11 are again of independent interest. We say that the contraction of an arbitrary (possibly not even connected) subset A of vertices to a single vertex is a k -contraction if there is a vertex v such that A is contained in the k -neighborhood of v . A k -contraction of a graph is obtained by simultaneously performing k -contractions of pairwise disjoint sets. Note that k -contractions can absorb local dense parts of a graph while at the same time approximately preserving distances. In particular, if a k -contraction of a graph G admits a weak r -neighborhood cover, then also G admits a weak r -neighborhood cover with the same degree and a slightly larger spread (depending on k). We prove that for every structurally nowhere dense class \mathcal{C} and every $G \in \mathcal{C}$ we can find an 8-contraction G' of G such that the class of all G' is almost nowhere dense. We will formalize the notion of almost nowhere denseness via the so-called *generalized coloring numbers*. We defer the formal definition of the *weak r -coloring number* of a graph G , denoted $\text{wcol}_r(G)$, to Section 7.2. As proved in [GKS17], every graph admits an r -neighborhood cover with spread $2r$ and degree $\text{wcol}_{2r}(G)$, hence, Theorem 11 will follow almost immediately from the following theorem.

Theorem 12. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $G \in \mathcal{C}$ there exists an 8-contraction $\text{contract}(G)$ of G , which is sparse in the following sense: for every $\varepsilon > 0$ and $r \in \mathbb{N}$ there exists $c(r, \varepsilon)$ such that for every $G \in \mathcal{C}$*

$$\text{wcol}_r(\text{contract}(G)) \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

2 Technical Overview

In this section, we give a more detailed technical overview of our proof. For this overview we assume some background from graph theory and logic. We will provide all formal definitions in Section 3 below.

As outlined above, the approach of Grohe, Kreutzer and Siebertz [GKS17] on nowhere dense classes is based on the locality properties of first-order logic. By Gaifman’s Locality Theorem [Gai82], the problem to decide whether a general first-order sentence φ is true in a graph can be reduced to testing whether other *local* formulas are true in the graph. To evaluate local formulas, we can restrict to bounded-radius neighborhoods of the graph, where the radius depends only on φ .

As proved in [GKS17], the r -neighborhoods of vertices from a graph from a nowhere dense class behave well, which is witnessed by a characterization of nowhere dense graph classes in terms of a game, called the *Splitter game*. In the radius- r Splitter game, two players called Connector and Splitter, engage on a graph and thereby recursively decompose local neighborhoods. Starting with the input graph G , in each of the following rounds, Connector chooses a subgraph of the current game graph of radius at most r and Splitter deletes a single vertex from this graph. The game continues with the resulting graph and terminates when the empty graph is reached. A class of graphs is nowhere dense if and only if for every r there exists ℓ such that Splitter can win the radius- r Splitter game in ℓ rounds.

While this game characterization shows that graphs from nowhere dense classes have simple neighborhoods, it does not immediately lead to an efficient model checking algorithm. There are two central challenges that need to be overcome:

- The first problem is that the algorithm needs to be called recursively on the local neighborhoods, where the graphs in the recursion get simpler and simpler by deleting vertices guided by the Splitter game. The deletion of a vertex can be encoded by coloring the neighbors of the vertex, so that the formulas can be rewritten to equivalent formulas, which, however, have to

be localized again in each recursive step. By simply applying Gaifman’s theorem again, this leads to an increase in the quantifier rank, and hence of the locality radius of the formulas, so that one can no longer play the Splitter game with the original radius. It is forbidden to increase the radius of the game during play. This problem was handled in [GKS17] by establishing a rank-preserving local normal form for first-order formulas, where localization is possible without increasing the quantifier rank.

- As mentioned before, a second problem arises as one cannot simply recursively branch into all local neighborhoods, as this would in the worst case lead to a branching degree of n , and a running time of $\Omega(n^\ell)$, where ℓ is the depth of the recursion. Therefore, nearby neighborhoods were clustered and handled together in one recursive call using sparse r -neighborhood covers. As proved in [GKS17] for every nowhere dense graph class \mathcal{C} , every r and $\varepsilon > 0$ there exists a constant $c(r, \varepsilon)$ such that every $G \in \mathcal{C}$ admits an r -neighborhood cover with spread at most $2r$ and degree at most $c(r, \varepsilon) \cdot |G|^\varepsilon$. With some technical work, formulas were incorporated into the neighborhood covers. Finally, by setting $\rho = \varepsilon/\ell$ and starting with an r -neighborhood cover of degree $c(r, \rho) \cdot |G|^\rho$, the complete recursive data structure could be constructed in time $c(r, \rho) \cdot |G|^{1+\varepsilon}$.

Very recently, based on a notion of *flip wideness* [DMST22], monadically stable graph classes were characterized by a game, called the *Flipper game* [GMM⁺23], which is similar in spirit to the Splitter game. In the radius- r Flipper game, two players called Connector and Flipper, engage on a graph and thereby recursively decompose local neighborhoods. Starting with the input graph G , in each of the following rounds, Connector chooses a subgraph of the current game graph of radius at most r and Flipper chooses two sets of vertices A and B and flips the adjacency between the vertices of these two sets. The game continues with the resulting graph. Flipper wins once a graph consisting of a single vertex is reached. A class of graphs is monadically stable if and only if for every r there exists $\ell(r)$ such that Flipper can win the radius- r Flipper game in $\ell(r)$ rounds. This new characterization suggests approaching the model checking problem just as on nowhere dense classes, where we just replace the Splitter game by the Flipper game. However, on dense graphs, both of the above challenges reveal additional difficulties:

- The by far most involved aspect of the algorithm in [GKS17] was the introduction of a rank-preserving local normal form to keep the quantifier rank steady during each localization step. Even more problems arise when adapting the rank-preserving normal form to account for flips instead of vertex deletions. For example, in the construction of [GKS17], to avoid introducing new quantifiers, some local distances have to be encoded by colors around deleted elements. More precisely, when a vertex v is deleted, one colors for all $i \leq k$ the vertices at distance i from v with a predicate D_i . Then one can query $\text{dist}(x, y) \leq d$ for $d \leq k$ in $G - v$ using the formula $\text{dist}(x, y) \leq d \vee \bigvee_{i+j \leq d} D_i(x) \wedge D_j(y)$. Note that this trick is not possible if we deal with a flip between A and B , since A and B may be arbitrarily large, and we would have to introduce a distance atom for each element of $A \cup B$. We completely sidestep this and other problems arising from the rank-preserving local normal form by instead building a much simpler and more powerful rank-preservation mechanism using *local types*, which we explain soon.
- For nowhere dense graph classes, the existence and algorithmic construction of sparse neighborhood covers follows elegantly from a bound on the so-called generalized coloring numbers. These arguments are specific to nowhere dense classes and do not transfer to more general classes. In the next few pages, we explain how to overcome this problem.

Outline of the Algorithm. Locality of first-order logic is a key tool both for the study of the expressive power of FO and for logic-based applications. For a graph G , $v \in V(G)$ and $q \in \mathbb{N}$, we call the set of all formulas $\varphi(x)$ of quantifier rank at most q such that $G \models \varphi(v)$ the q -type of v in G . We write $N_r[v]$ for the closed r -neighborhood of v in G . By Gaifman’s Locality Theorem [Gai82], there is a function $g: \mathbb{N} \rightarrow \mathbb{N}$ such that the following holds. Whenever two vertices $u, v \in V(G)$ have the same $g(q)$ -type in $G[N_{7^q}[u]]$ and $G[N_{7^q}[v]]$, respectively, then they have the same q -type in G . This statement is made explicit, for example, in the proof of Lemma 1.5.2 of [EF99].

One obvious reason for the increase of quantifier rank from q to $g(q)$ that happens in the proof of Gaifman’s theorem is that new quantifiers are needed to express distances. This can be avoided by introducing distance atoms, which led to the technical rank-preserving normal form for FO in [GKS17]. Another not so obvious reason for the increase of quantifier rank is that nearby elements whose local neighborhoods overlap must be handled in a non-trivial combinatorial way, which in Gaifman’s theorem leads to a locality radius of 7^q . This is not satisfying because it is well known that FO with q quantifiers can express distances only up to 2^q .

This observation led to the notion of *local types*, which were introduced in [GPPT22] in the context of twin-width. The *localization* of a formula φ with free variables is the formula with the same free variables as φ that replaces every subformula $\exists x \psi(x, \bar{y})$ with quantifier rank k with $\exists x \in N_{2^{k-1}}[\bar{y}] \psi(x, \bar{y})$. Likewise, every subformula $\forall x \psi(x, \bar{y})$ with quantifier rank k is replaced with $\forall x \in N_{2^{k-1}}[\bar{y}] \psi(x, \bar{y})$. We call a formula *local* if it is the localization of some formula. The *local q -type* of $v \in V(G)$ is the set of all local formulas $\varphi(x)$ such that $G \models \varphi(v)$. Observe that syntactic localization does not increase the quantifier rank of a formula, since ψ has quantifier rank $k - 1$ and with this number of quantifiers we can express distances up to 2^{k-1} .

We follow the approach of [GPPT22] to relate local types and a local variant of Ehrenfeucht-Fraïssé games and extend the results of [GPPT22] to provide a fine-grained analysis of the locality properties of first-order logic. We prove (Theorem 4) that for any two sets $U_1, U_2 \subseteq V(G)$ whose local neighborhoods look alike and which are sufficiently far from each other, if we find an element $a \in U_1$ satisfying a first-order property $\varphi(x)$, then we also find an element $b \in U_2$ satisfying the same property.

We now incorporate ideas from an elegant approach to model checking of [GGK20], showing that FO model checking can be reduced to deciding whether two vertices have the same q -type. In this approach, the recursive evaluation trees are reduced by keeping only a bounded number of representative vertices. Our technical implementation of this idea is independent from [GGK20], and instead based on *guarded formulas*, which are naturally defined as follows. Given a set of unary predicates \mathcal{U} , we say that a formula is \mathcal{U} -*guarded* if every quantifier is of the form $\exists x \in U$ or $\forall x \in U$ for some $U \in \mathcal{U}$. It is an easy observation that when evaluating guarded sentences, we can ignore all vertices outside the guarding sets. We now proceed to compute a constant number t of representative guarding sets U_1, \dots, U_t , where t depends only on φ and the class \mathcal{C} . While the U_i can contain arbitrarily many vertices (here we also deviate from the approach of [GGK20]), each U_i is contained in a local neighborhood of the input graph. The candidate sets for U_i are derived from the recursive computation of types of clusters from a sparse weak neighborhood cover, where the recursion is guided by the Flipper game. As we proved before, for two sets $U_1, U_2 \subseteq V(G)$ whose local neighborhoods look alike and which are far from each other, if we find an element $a \in U_1$ satisfying a first-order property $\varphi(x)$, then we also find an element $b \in U_2$ satisfying the same property, which allows us to reduce to a bounded number of representative sets. By appropriately grouping the remaining U_i , we have reduced the model checking problem to a local problem, where the locality radius remains stable over the recursion.

Our approach avoids the complicated construction of the rank-preserving local normal form in the model checking algorithm of [GKS17]. It also greatly simplifies the interplay between local formulas and neighborhood covers. We believe that our results are of interest beyond the scope of this paper. **The idea of using local types in combination with neighborhood covers for model checking in sparse graphs arose in discussions with Szymon Toruńczyk. This approach is also explored in our companion paper with Szymon, where we present a simplified proof of model checking on nowhere dense classes [AT].** We want to thank Szymon for these many useful discussions.

Contribution: Construction of Neighborhood Covers. The missing piece for our model checking algorithm is the construction of sparse neighborhood covers. As a second main contribution of our paper, we prove in Theorem 10 that weak r -neighborhood covers on general graphs can be efficiently approximated.

To prove the theorem, we provide a robust ILP formulation for weak r -neighborhood covers. A solution to the LP relaxation can then be turned efficiently into a *fractional weak r -neighborhood cover*, which is a set of covering clusters equipped with a real value between 0 and 1, which can intuitively be understood as a probability assignment, such that the sum of values for each neighborhood to be covered sum up to at least one. Crucially, the constructed fractional weak r -neighborhood covers have at most $\mathcal{O}(n^2)$ clusters. Now, instead of having to consider the exponential number of all subsets of s -neighborhoods, we can search for our r -neighborhood cover among the at most $\mathcal{O}(n^2)$ clusters. Via randomized rounding, we turn the fractional weak r -neighborhood cover into a weak r -neighborhood cover whose degree is at most a logarithmic factor larger. This algorithm can be derandomized using standard methods. The running time of this procedure, and in fact of our whole model checking algorithm, is dominated by the time needed to solve LPs.

We want to highlight that our algorithm based on ILP formulations and randomized rounding is elementary and easy to understand. This is in sharp contrast to the previous algorithm on nowhere dense graph classes, which relied on the structure theory for these graph classes.

Contribution: Structurally Nowhere Dense Graph Classes. Finally, we demonstrate that structurally nowhere dense graph classes admit flip-closed sparse weak neighborhood covers and thereby establish fixed-parameter tractability of FO model checking on these classes, that is, Theorem 1.

To prove the existence of sparse weak neighborhood covers on structurally nowhere dense graph classes, we delve into the structure theory for these graph classes and again develop tools that we believe are interesting beyond the scope of the paper. To describe our contribution, let us first give an intuitive overview over nowhere dense and structurally nowhere dense classes. Nowhere dense graph classes are very general classes of sparse graphs. They were originally defined as classes \mathcal{C} such that for every radius $r \in \mathbb{N}$, some graph H_r is excluded as a *depth- r minor* in graphs from \mathcal{C} . Intuitively, a depth- r minor of a graph is obtained by contracting pairwise disjoint connected subgraphs of radius at most r to single vertices. Observe that the class property of nowhere denseness is preserved under taking depth- r minors for any fixed r . Nowhere dense graph classes are very well understood and admit many combinatorial characterizations. In particular, nowhere dense graph classes admit sparse neighborhood covers. As mentioned before, the construction of sparse neighborhood covers for these classes is based on a characterization via generalized coloring numbers $\text{wcol}_r(G)$. As proved in [GKS17], every graph admits an r -neighborhood cover with spread $2r$ and degree $\text{wcol}_{2r}(G)$.

For our construction of neighborhood covers we work with local contractions instead of bounded depth minors. Observe that k -contractions do not necessarily preserve the class property of nowhere denseness, however, they preserve distances up to factors $(2k + 1)$, and hence the property of admitting sparse weak neighborhood covers. In particular, neighborhood covers can be lifted from a k -contraction H of a graph G to the original graph G with the same degree and a spread that is larger by at most a factor of $(2k + 1)$. The use of the k -contraction is to absorb local dense parts of the graph into single vertices. As a key result, we obtain the following theorem.

Theorem 12. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $G \in \mathcal{C}$ there exists an 8-contraction $\text{contract}(G)$ of G , which is sparse in the following sense: for every $\varepsilon > 0$ and $r \in \mathbb{N}$ there exists $c(r, \varepsilon)$ such that for every $G \in \mathcal{C}$*

$$\text{wcol}_r(\text{contract}(G)) \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

Note that the theorem is only existential, and we cannot efficiently construct the 8-contraction $\text{contract}(G)$ for a given graph G . Nevertheless, the theorem gives strong insights into the structure of graphs from structurally nowhere dense classes, and as described above, we can then derive the existence of sparse weak r -neighborhood covers for graphs from these classes. Since the property of structural nowhere denseness is preserved under flips, and from Theorem 12 one can derive the existence of weak neighborhood covers with a spread depending only on r and not on the class, we naturally derive the existence of flip-closed sparse weak neighborhood covers for structurally nowhere dense classes, that is, Theorem 11.

The key to proving Theorem 12 is based on a structural characterization of structurally nowhere dense graph classes in terms of *quasi-bushes* [DGK⁺22a]. Quasi-bushes have the form of a tree of bounded depth with additional links that define the edges of the decomposed graph. Furthermore, the quasi-bushes of structurally nowhere dense classes have bounded weak r -coloring numbers. The sought k -contractions will be found by careful contraction of certain subtrees of the quasi-bushes.

3 Preliminaries

We write \mathbb{N} for the set of natural numbers $\{1, 2, \dots\}$. For $m \in \mathbb{N}$ we let $[m] = \{1, \dots, m\}$. We write \bar{x}, \bar{y}, \dots for tuples of variables and $\bar{a}, \bar{b}, \bar{v}, \bar{w}, \dots$ for tuples of elements and usually leave it to the context to determine the length of a tuple. We access the elements of a tuple using subscripts, that is, $\bar{x} = x_1 x_2 \dots x_{|\bar{x}|}$.

3.1 Graphs

All graphs in this paper are finite, loopless, and vertex-colored. More precisely, a graph G is a relational structure with a finite universe $V(G)$ over a finite signature Σ consisting of the binary, irreflexive edge relation E and a finite number of unary color predicates. Note that we do not require that colors from Σ are interpreted by disjoint sets, that is, vertices may carry multiple colors. Most of the time, the signature Σ will be clear from the context and we will not mention it explicitly. We will commonly expand graphs with additional colors. For a graph G over the signature Σ and a subset of its vertices $W \subseteq V(G)$, we write $G\langle X \mapsto W \rangle$ for the graph G over the signature $\Sigma \cup \{X\}$ where X is interpreted as W , that is, in G we color the vertices of W with the new color X . We write $G\langle W \rangle$ as a shorthand for $G\langle W \mapsto W \rangle$, where by slight abuse of notation we identify a relation symbol with its interpretation. For a family $\mathcal{U} = \{U_1, \dots, U_t\}$ of subsets of $V(G)$, we write $G\langle \mathcal{U} \rangle$ or $G\langle U_1, \dots, U_t \rangle$ as a shorthand for $G\langle U_1 \rangle \dots \langle U_t \rangle$.

Unless explicitly stated otherwise, graphs are undirected, that is, we assume that E is interpreted by a symmetric and irreflexive relation. We often denote an undirected edge between u and v by $\{u, v\}$ and a directed edge, which we will call an *arc*, pointing from u to v by (u, v) . We write $E(G)$ for the edge set of a graph G and $|G|$ for the number of its vertices. We use the standard notation from graph theory from Diestel's textbook [Die12], which we extend to colored graphs in the natural way.

Induced subgraphs. For a set of vertices $X \subseteq V(G)$, we write $G[X]$ for the subgraph of G induced by X , and $G - X$ for the subgraph of G induced by $V(G) - X$. We say a class of graphs is *hereditary* if it is closed under taking induced subgraphs.

Distances and neighborhoods. Let G be a graph. For two vertices $u, v \in V(G)$, we write $\text{dist}(u, v)$ for the distance between u and v , which is set to ∞ if u and v are not connected in G . If \bar{a} and \bar{b} are tuples (or sets) of vertices we write $\text{dist}(\bar{a}, \bar{b})$ for the minimum distance between some $a \in \bar{a}$ and some $b \in \bar{b}$. For $r \in \mathbb{N}$ and $v \in V(G)$ we write $N_r[v] := \{u \in V(G) \mid \text{dist}(u, v) \leq r\}$ for the (closed) r -neighborhood of v in G , and more generally, for a tuple (or set) \bar{a} we let $N_r[\bar{a}] = \bigcup_{a \in \bar{a}} N_r[a]$.

3.2 Logic

We use standard terminology from model theory and refer to [Hod97] for extensive background. Every formula in this paper will be a first-order formula. However, we will often not explicitly write down the formulas if the properties they express are obviously expressible. For example, $x \in N_r[\bar{y}]$ stands for the first-order formula expressing that x is contained in the r -neighborhood of \bar{y} . Also for a color predicate P , we often write $\exists x \in P \varphi$ as a shorthand for $\exists x P(x) \wedge \varphi$ and $\forall x \in P \varphi$ as a shorthand for $\forall x P(x) \rightarrow \varphi$. For a formula φ , we write $\text{free}(\varphi)$ for the set of free variables appearing in φ , and we write $\varphi(\bar{x})$ to indicate that the free variables of φ are in \bar{x} .

Every formula $\eta(x, y)$ on a graph G defines the relation $\eta(G) := \{(u, v) \in V(G)^2 \mid G \models \eta(u, v)\}$. Similarly, a formula $\nu(x)$ defines the set $\{v \in V(G) \mid G \models \nu(v)\}$. We call a formula $\eta(x, y)$ symmetric and irreflexive if on all graphs the relation it defines is symmetric and irreflexive.

Normalization. For every finite signature Σ , quantifier rank q , and tuple of free variables \bar{x} , up to equivalence there only exist a finite number of distinct formulas $\varphi(\bar{x})$ over Σ with quantifier rank at most q . Testing equivalence of first-order formulas is undecidable. However, given a formula we can effectively compute an equivalent *normalized* formula of the same quantifier rank, such that again there only exist a finite number of distinct normalized formulas $\varphi(\bar{x})$ over Σ with quantifier rank at most q . In particular, the length of a normalized formula $\varphi(\bar{x})$ with quantifier rank q over Σ only depends on $|\bar{x}|$, q , and Σ . The normalization process works by renaming quantified variables, reordering boolean combinations into conjunctive normal form, and deleting duplicates from conjunctions and disjunctions. We will assume throughout this paper that all appearing formulas are normalized. This also includes formulas which we construct ourselves: normalization is always performed implicitly as the last step of a construction.

Types. Let G be a graph and $\bar{a} \in V(G)^{|\bar{a}|}$ be a tuple in G . We denote by $\text{tp}_q(G, \bar{a})$ the finite set of all normalized formulas $\varphi(\bar{x})$ with $|\bar{x}| = |\bar{a}|$ and quantifier rank at most q over the signature of G such that $G \models \varphi(\bar{a})$. We write $\text{tp}_q(G) := \text{tp}_q(G, \emptyset)$ for the set of all normalized sentences of quantifier rank at most q that hold in G .

3.3 Stability

The bipartite graph $(\{a_1, \dots, a_\ell\}, \{b_1, \dots, b_\ell\}, E)$, where $a_i b_j \in E(G)$ if and only if $i \leq j$, is called the half graph of order ℓ . The half graph of order 5 is depicted in Figure 1.

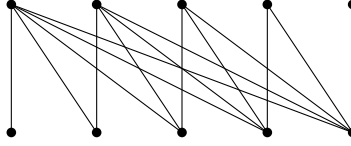


Figure 1: The half graph of order 5.

Half graphs form the graph theoretical equivalent of linear orders. Intuitively, *monadically stable* classes of graphs are those classes which do not admit the encoding of arbitrary large half graphs. The precise notion of encoding will be that of a *transduction*. A transduction is an operation mapping an input graph with signature Σ to a set of output graphs with the same signature². Every transduction T consists of two first-order formulas $\nu(x)$ and $\eta(x, y)$ over the signature $\Sigma \uplus \Gamma$ for an infinite set of color predicates Γ , where η is symmetric and irreflexive. The set of output graphs $T(G)$ is generated from the input graph G as follows.

1. G is mapped to its (infinite) set of Γ -colorings.
2. Every colored graph G^+ is mapped to a graph H with signature Σ ,
 - whose vertex set $\{v \in V(G^+) \mid G^+ \models \nu(v)\}$ is defined by ν ,
 - whose edge set $\{\{u, v\} \mid G^+ \models \eta(u, v), u, v \in V(H)\}$ is defined by η , and
 - where the color predicates from Σ are interpreted as in G .

As η and ν can only reference a finite number of colors, the resulting set of output graphs is finite as well. As η is symmetric and irreflexive, every output graph is undirected and loopless.

The definition of transductions lifts to classes of graphs where we say a class \mathcal{D} is a *transduction* of a class \mathcal{C} if there exists a fixed transduction T that produces from \mathcal{D} all graphs from \mathcal{C} , that is, $\mathcal{D} \subseteq \bigcup_{G \in \mathcal{C}} T(G)$. A class \mathcal{D} is called *structurally nowhere dense* if it is a transduction of a nowhere dense class \mathcal{C} . The notion of monadic stability is originally defined in terms of more general interpretations. However, by results of [BS85] we can define them in terms of transductions. We call a class \mathcal{C} *monadically stable* if the class of all half graphs cannot be transduced in \mathcal{C} . Every structurally nowhere dense class is also monadically stable [PZ78, AA14]. Generalizing this notion, we call a class \mathcal{C} *monadically NIP* if the class of all graphs cannot be transduced in \mathcal{C} .

Transductions can be composed, hence, for classes $\mathcal{C}, \mathcal{D}, \mathcal{E}$ we have that if \mathcal{E} is a transduction of \mathcal{D} and \mathcal{D} is a transduction of \mathcal{C} , then also \mathcal{E} is a transduction of \mathcal{C} . Combining this observation with the fact that there exists a transduction yielding all induced subgraphs of an input graph, we will from now on assume without loss of generality that all structurally nowhere dense and monadically stable classes are hereditary.

²Transductions may also be defined with arbitrary (and differing) input and output signatures and may include copy operations. However, we will not need this flexibility.

3.4 Flipper Game

A *flip* $F = (A, B)$ is a pair of sets of vertices. We write $G \oplus F$ for the graph on the vertex set of G where the adjacency between vertices of A and B is complemented, that is, we have $\{u, v\} \in E(G \oplus F)$ if and only if $\{u, v\} \in E(G)$ xor $(u, v) \in (A \times B) \cup (B \times A)$. With this definition, we have $G \oplus (A, B) = G \oplus (A \cap V(G), B \cap V(G))$, so there is no need to require A, B to be subsets of $V(G)$. This is useful for working with induced subgraphs. For a set $F = \{F_1, \dots, F_n\}$ of flips, we write $G \oplus F$ for the graph $G \oplus F_1 \oplus \dots \oplus F_n$. Note that the order in which we carry out the flips does not matter.

We are going to use the *Flipper game*, which was introduced in [GMM⁺23] and is defined as follows.

Definition 1 (Flipper game). Fix a radius r . The *radius- r Flipper game* is played by two players, *Flipper* and *Connector*, on a graph G as follows. At the beginning, set $G_0 := G$. In the i th round, for $i > 0$, the game proceeds as follows.

- If $|G_{i-1}| = 1$, then Flipper wins.
- Connector chooses G_{i-1}^{loc} as the subgraph of G_{i-1} induced by a subset³ of an r -neighborhood in G_{i-1} .
- Flipper chooses a flip F and applies it to produce G_i , that is, $G_i = G_{i-1}^{\text{loc}} \oplus F$.

We will use the fact that, when playing in monadically stable classes, Flipper has an efficient, algorithmic strategy that wins the Flipper game in a bounded number of rounds. We will formalize this fact by defining *strategies* and their corresponding *runtime*.

Strategies. Fix a graph class \mathcal{C} and a radius $r \in \mathbb{N}$. A *radius- r Connector strategy* is a function

$$\text{con} : (G_i) \mapsto (G_i^{\text{loc}})$$

mapping the graph G_i of round i to the graph G_i^{loc} induced by a subset of an r -neighborhood in G_i . A *radius- r Flipper strategy* is a function

$$\text{flip} : (G_i^{\text{loc}}, \mathcal{I}_i) \mapsto (F, \mathcal{I}_{i+1})$$

mapping Connectors move G_{i-1}^{loc} of round i to a flip F . Additionally, Flipper receives an internal state \mathcal{I}_i that is updated to a successor state \mathcal{I}_{i+1} during the computation. In the algorithmic context, this is a convenient abstraction from the usual definition of strategies based on game histories. We will not define the precise shape of an internal state as it is an implementation-detail which may vary between different Flipper strategies. Flipper can utilize the internal states, for example, as a memory to store past Connector moves or precomputed flips for future turns. An initial state $\mathcal{I}_0 = \mathcal{I}_0(\text{flip}, G)$ will be computed from the initial graph at the beginning of the Flipper game.

Given radius- r Connector and Flipper strategies con and flip and a graph $G \in \mathcal{C}$, the *Flipper run* $\mathcal{R}(\text{con}, \text{flip}, G)$ is the infinite sequence of *positions*

$$\mathcal{R}(\text{con}, \text{flip}, G) = (G_0, \mathcal{I}_0), (G_1, \mathcal{I}_1), (G_2, \mathcal{I}_2), (G_3, \mathcal{I}_3), \dots$$

³In the definition of the Flipper game given in [GMM⁺23], Connector is required to choose the entire r -neighborhood of v as G_{i-1}^{loc} . The version defined in this paper is referred to as *Induced-Subgraph-Flipper game* in [GMM⁺23]. However, the difference is insignificant. Using the algorithmic strategy provided in [GMM⁺23], Flipper wins both versions of the game in monadically stable classes. For brevity we will always refer to the Induced-Subgraph-Flipper game as Flipper game.

such that $G_0 = G$ and $\mathcal{I}_0 = \mathcal{I}_0(\text{flip}, G)$, and for all $i \geq 0$ we have

$$(G_{i+1} = G_i^{\text{loc}} \oplus F, \mathcal{I}_{i+1}) \text{ where } G_i^{\text{loc}} = \text{con}(G_i) \text{ and } (\mathcal{I}_{i+1}, F) = \text{flip}(\mathcal{I}_i, G_i^{\text{loc}}).$$

A *winning position* is a tuple (G_i, \mathcal{I}_i) such that G_i contains only a single vertex. A radius- r Flipper strategy flip is ℓ -*winning*, if for every $G \in \mathcal{C}$ and for every radius- r Connector strategy con , the ℓ th position of $\mathcal{R}(\text{con}, \text{flip}, G)$ is a winning position. Note that, while $\mathcal{R}(\text{con}, \text{flip}, G)$ is an infinite sequence, once a winning position is reached, it is only followed by winning positions.

Runtimes. Let \mathcal{C} be a class of graphs, let $r \in \mathbb{N}$, and let flip be a radius- r Flipper strategy. For a graph $G \in \mathcal{C}$ and a radius- r Connector strategy con , let $t_0(\text{con}, G)$ be the time needed to compute the initial internal state $\mathcal{I}_1(\text{flip}, G)$. For $i \geq 0$, let $t_{i+1}(\text{con}, G)$ be the time needed to compute the output of flip given as input $\text{con}(G_i)$ and \mathcal{I}_i , where (G_i, \mathcal{I}_i) is the $(i+1)$ th position in $\mathcal{R}(\text{con}, \text{flip}, G)$. The *runtime* of flip on \mathcal{C} is the function f defined by

$$f(n) = \max_{G \in \mathcal{C}, |G| \leq n} \sup_{\substack{\text{radius-}r \text{ Connector} \\ \text{strategy } \text{con}}} \sup_{i \geq 0} t_i(\text{con}, G).$$

In words, $f(n)$ is the maximum time need by Flipper to compute a move on any play of the game on a graph from the class \mathcal{C} .

We are now ready to state the following theorem, which is one of the main results of [GMM⁺23].

Theorem 2 ([GMM⁺23, Theorem 11.2]). *For every monadically stable class \mathcal{C} there exists a function $f : \mathbb{N} \rightarrow \mathbb{N}$, such that for every radius $r \in \mathbb{N}$ there exist $\ell \in \mathbb{N}$ and an ℓ -winning radius- r Flipper strategy flip^* with runtime $f(r) \cdot n^2$.*

For every monadically stable class \mathcal{C} and radius $r \in \mathbb{N}$, the algorithmic strategy flip^* given by Theorem 2 is the one we will be using throughout the paper. We define $\text{game-depth}(\mathcal{C}, r)$ to be the bound on the number of rounds needed for Flipper to win the radius- r Flipper game on any graph from \mathcal{C} while following flip^* . We will call a sequence of positions $\mathcal{H} = (G_0, \mathcal{I}_0), \dots, (G_\ell, \mathcal{I}_\ell)$ a (\mathcal{C}, r) -*history of length ℓ* , if it is a prefix of the Flipper run $\mathcal{R}(\text{con}, \text{flip}^*, G_0)$ for some radius- r Connector strategy con and some graph $G_0 \in \mathcal{C}$. Note that, by definition, the time needed to calculate Flipper's next move $\text{flip}^*(G_\ell^{\text{loc}} = \text{con}(G_\ell), \mathcal{I}_\ell)$ depends on the size of G_0 and not on the (possibly much smaller) size of G_ℓ . Also note that, in order to apply flip^* , we only require G_0 to be contained in \mathcal{C} . The current input graph G_ℓ^{loc} might not be contained in \mathcal{C} and, indeed, this is usually the case since G_ℓ^{loc} was obtained from G_0 by ℓ rounds of flipping (and localizing). However as localizing and flipping for a bounded number of rounds is expressible by a transduction, if \mathcal{C} is monadically stable (structurally nowhere dense), then G_ℓ^{loc} will be from a class that is still monadically stable (structurally nowhere dense).

3.5 Weak Neighborhood Covers

A *weak r -neighborhood cover* with *degree d* and *spread s* of a graph G is a family \mathcal{X} of subsets of $V(G)$, called *clusters*, such that

- every cluster is a subset of an s -neighborhood in G , that is, for every $X \in \mathcal{X}$ there exists a vertex $\text{center}(X) \in V(G)$ with $X \subseteq N_s[\text{center}(X)]$,
- the r -neighborhood of every vertex is contained in some cluster, that is, for every $w \in V(G)$ there exists cluster $\text{cluster}(w) \in \mathcal{X}$ with $N_r[w] \subseteq \text{cluster}(w)$, and
- every vertex occurs in at most d clusters, that is, for all $v \in V(G)$ we have

$$|\{X \in \mathcal{X} \mid v \in X\}| \leq d.$$

We note that the above definition is a relaxation of the more common notion of an (non-weak) *r -neighborhood cover*, where for every cluster X , the induced subgraph $G[X]$ is required to have radius at most s .

Definition 2. A graph class \mathcal{C} admits *sparse* weak neighborhood covers if there exist functions $g(r, \varepsilon)$ and $\sigma(r) \geq r$ such that for every $r \in \mathbb{N}$, every $\varepsilon > 0$, every n -vertex graph $G \in \mathcal{C}$ admits a weak r -neighborhood cover with degree $g(r, \varepsilon) \cdot n^\varepsilon$ and spread $\sigma(r)$.

When \mathcal{C} is a class of graphs and $\ell \in \mathbb{N}$, we write \mathcal{C}_ℓ for the class containing all graphs that can be obtained by applying at most ℓ flips to graphs from \mathcal{C} . The following definition will be the key to decompose graphs into their local neighborhoods during a play of the Flipper game.

Definition 3. A class \mathcal{C} of graphs admits *flip-closed* sparse weak neighborhood covers if there exist functions $g(r, \varepsilon, \ell)$ and $\sigma(r) \geq r$ such that for every $r \in \mathbb{N}$, every $\varepsilon > 0$, and every $\ell \in \mathbb{N}$, every n -vertex graph $G \in \mathcal{C}_\ell$ admits a weak r -neighborhood cover with degree $g(r, \varepsilon, \ell) \cdot n^\varepsilon$ and spread $\sigma(r)$.

In the above definition, it is crucial that the spread bound σ remains independent of ℓ .

4 Guarded Formulas and Local Types

4.1 Guarded Formulas

Given a set of unary predicates \mathcal{U} , we say a formula is *\mathcal{U} -guarded* if every quantifier is of the form $\exists x \in U$ or $\forall x \in U$ for some $U \in \mathcal{U}$. Our model checking algorithm crucially builds on the simple observation that when evaluating guarded sentences, we can ignore all vertices outside the guarding sets.

Observation 1. Given a graph G and a family $\mathcal{U} = \{U_1, \dots, U_t\}$ of subsets of $V(G)$. Interpreting each set from $U \in \mathcal{U}$ as a unary predicate, we have for every \mathcal{U} -guarded sentence φ that

$$G \langle U_1, \dots, U_t \rangle \models \varphi \iff G \langle U_1, \dots, U_t \rangle [U_1 \cup \dots \cup U_t] \models \varphi.$$

Our goal is to compute a representative set of guards $\mathcal{U} = \{U_1, \dots, U_t\}$ such that we can translate our input formula φ into an equivalent \mathcal{U} -guarded formula. Here, crucially, the size t of \mathcal{U} shall depend only on φ (and eventually the depth of the recursion on \mathcal{C}). Assume for now that we have recursively computed a large set of candidate guards $\{V_1, \dots, V_m\}$. Then the selection of the set \mathcal{U} is based on the following key theorem that we prove in the remainder of this section.

Theorem 4. *Let G be a graph and let $A, B \subseteq V(G)$ be vertex sets such that $\text{dist}(A, B) > 2^k$ and $\text{tp}_k(G \langle X \mapsto A \rangle [N_{2^{k-1}-1}[A]]) = \text{tp}_k(G \langle X \mapsto B \rangle [N_{2^{k-1}-1}[B]])$. Let $\bar{w} \in V(G)^{|\bar{y}|}$ be vertices with $\text{dist}(\bar{w}, A \cup B) \geq 2^k$. Then for every formula $\varphi(\bar{y}, x)$ of quantifier rank at most $k-1$ in the signature of G we have $G \langle A \rangle \models \exists x \in A \varphi(\bar{w}, x) \iff G \langle B \rangle \models \exists x \in B \varphi(\bar{w}, x)$.*

Intuitively, the theorem states the following. Given a graph G and two sets A and B whose neighborhoods look alike and which are far from each other. If we find an element $a \in A$ satisfying a first-order property $\varphi(x)$, then we also find an element $b \in B$ satisfying the same property, which will allow us to restrict quantification to appropriately chosen guard sets.

Note that the locality radius in the theorem naturally corresponds to distances that can be expressed with $k-1$ quantifiers. The proof of the theorem is based on the notion of *local types*, which were introduced in [GPPT22]. Local types over graphs beautifully capture the locality properties of FO by identifying the semantic restriction to 2^{k-1} -neighborhoods with the ability of FO to syntactically make these restrictions. We remark that the results of this section do *not* directly translate to structures with relations of arity greater than 2, since defining distances in (the Gaifman graph of) such structures may require the use of additional quantifiers. Some of the results we prove here were proved in a different notation already in [GPPT22] and in the lecture notes of Szymon Toruńczyk [Tor22], while some lemmas and in particular the main theorem of this section, Theorem 4, is new. We provide all proofs for consistency and completeness.

Our proof of Theorem 4 proceeds as follows. In Section 4.2 we recall the notion of Ehrenfeucht-Fraïssé games (short EF-games), which are a classical tool of (finite) model theory to understand the expressiveness of first-order logic. We introduce a local variant of the games in Section 4.3, where all moves are restricted to the local neighborhoods of elements that were played before. Classically, EF-games can be played on two different structures. In Section 4.4 we show that when playing on the same graph local games determine global games. We relate local games with local types in Section 4.5. Up to this point, most of the results were provided in a similar form already in [GPPT22]. Towards the proof of Theorem 4 we now extend the framework and incorporate guards into local games in Section 4.6.

4.2 Games

The EF-game is played by two players called *Spoiler* and *Duplicator* on two structures. It is Spoiler's goal to distinguish the two structures, while Duplicator wants to show that the structures cannot be distinguished. The connection with first-order logic is as follows: Duplicator has a winning strategy in the q -round EF-game on two structures if and only if the two structures satisfy the same sentences of quantifier rank at most q . In this work, we consider only games that are played on a single graph with different distinguished vertices \bar{a} and \bar{b} . We refer to the literature for extensive background on EF-games, for example, to the textbook [Lib04].

Each position of the game is a tuple (G, \bar{a}, \bar{b}, k) consisting of a graph G that is fixed throughout the game, two non-empty tuples of vertices \bar{a}, \bar{b} of equal length, and a counter $k \in \mathbb{N}$ that keeps track of the number of rounds that are still to play. The game starts in some position $(G, \bar{a}_0, \bar{b}_0, q)$. If we are currently at a position (G, \bar{a}, \bar{b}, k) , one round of the game proceeds as follows.

- Spoiler selects a vertex of G as a_k (he makes an *a-move*) or as b_k (he makes a *b-move*).
- If Spoiler made an *a-move*, then Duplicator has to reply with a *b-move*, that is, select a vertex of G as b_k , or if he made a *b-move*, then she has to reply with an *a-move*, that is, select a vertex of G as a_k .
- The game continues at position $(G, \bar{a}a_k, \bar{b}b_k, k-1)$.

The game terminates when $k = 0$. Assume that a final position $(G, a_\ell, \dots, a_1, b_\ell, \dots, b_1, 0)$ is reached ($\ell = k + |\bar{a}|$). We say this is a *winning position for Duplicator* if $(a_\ell, \dots, a_1, b_\ell, \dots, b_1)$ defines a *partial automorphism* on G , that is, for all $1 \leq i, j \leq \ell$,

- $a_i = a_j \iff b_i = b_j$,
- a_i and b_i have the same colors in G , and
- $(a_i, b_j) \in E(G) \iff (a_i, b_j) \in E(G)$.

We say that Duplicator has a *winning strategy* from a position if she can play such that she reaches – no matter how Spoiler plays – a winning position for Duplicator. Otherwise, we say Spoiler has a *winning strategy*. We write $(G, \bar{a}) \cong_k (G, \bar{b})$ if Duplicator has a winning strategy from position (G, \bar{a}, \bar{b}, k) . The proof of the following classical result can be found, for example, in [Lib04, Theorem 3.9].

Lemma 1. $(G, \bar{a}) \cong_k (G, \bar{b})$ if and only if $\text{tp}_k(G, \bar{a}) = \text{tp}_k(G, \bar{b})$.

4.3 Local Games

It is well known that first-order logic can express only local properties of graphs. In particular, for every k and $d \leq 2^k$ there exists a formula of quantifier rank k that can determine if the distance between two elements is exactly d , while there is no formula with k quantifiers that can distinguish between distances strictly greater than 2^k . This fact motivates our next definition of *local games*. The key observation is that in a position (G, \bar{a}, \bar{b}, k) when an element a_k at distance at most 2^{k-1} from \bar{a} is chosen by Spoiler, then Duplicator must respond with an element b_k at the exactly same distance to \bar{b} , (and vice versa) as otherwise Spoiler can change his strategy to simply point out the difference in distances. On the other hand, these locality properties imply that Spoiler will never select an element at distance greater than 2^{k-1} from both \bar{a} and \bar{b} , as this element could simply be copied by Duplicator.

At a position (G, \bar{a}, \bar{b}, k) , we say that a move (by Spoiler or Duplicator) is *local* if it is an a -move and contained in $N_{2^{k-1}}[\bar{a}]$ or if it is a b -move and contained in $N_{2^{k-1}}[\bar{b}]$. We define the *local EF-game* as the EF-game where we require that both players are only allowed to play local moves. We call the regular EF-game *global* to distinguish it from the local game. We write $(G, \bar{a}) \cong_k^{\text{local}} (G, \bar{b})$ if Duplicator has a winning strategy for the local game from position (G, \bar{a}, \bar{b}, k) .

4.4 Local Games Determine Global Games

We will argue that the local and global EF-games are equivalent when we start from positions (G, \bar{a}, \bar{b}, k) such that \bar{a} and \bar{b} are at distance greater than 2^{k+1} . Towards this goal, we formally prove the above observations. First, we observe that Duplicator has to respond to a local move of Spoiler with her own local move.

Lemma 2 (see Lemma 9.2 of [Tor22]). *Consider the global game at a position (G, \bar{a}, \bar{b}, k) . Assume Spoiler made a local a -move $a_k \in N_{2^{k-1}}[\bar{a}]$, say $a_k \in N_{2^{k-1}}[a_j]$ for $a_j \in \bar{a}$ and Duplicator answers with a b -move $b_k \notin N_{2^{k-1}}[\bar{b}]$, or symmetrically, Spoiler made a local b -move $b_k \in N_{2^{k-1}}[\bar{b}]$, say $b_k \in N_{2^{k-1}}[b_j]$ for some $b_j \in \bar{b}$ and Duplicator answers with an a -move $a_k \notin N_{2^{k-1}}[\bar{a}]$. Then Spoiler has a winning strategy for the remaining global game from position $(G, \bar{a}a_k, \bar{b}b_k, k-1)$. In particular, if Duplicator answers with a non-local move to a local move, she loses the game.*

Proof. We prove the statement by induction on k . By symmetry, we may assume that Spoiler makes an a -move. For $k = 1$ the claim is true, as $a_1 \in N_1[a_j] \Leftrightarrow b_1 \in N_1[b_j]$ is necessary for $(\bar{a}a_1, \bar{b}b_1)$ to be a partial automorphism.

Now assume $k > 1$. As $a_k \in N_{2^{k-1}}[a_j]$, Spoiler can play $a_{k-1} \in N_{2^{k-2}}[a_k] \cap N_{2^{k-2}}[a_j]$ as his next move. As $b_k \notin N_{2^{k-1}}[b_j]$, no matter which b_{k-1} Duplicator plays as a response, either $b_{k-1} \notin N_{2^{k-2}}[b_k]$ or $b_{k-1} \notin N_{2^{k-2}}[b_j]$. If $b_{k-1} \notin N_{2^{k-2}}[b_k]$, then by induction hypothesis applied to position $(G, a_k, b_k, k-1)$, Spoiler wins from position $(G, a_k a_{k-1}, b_k b_{k-1}, k-2)$. If $b_{k-1} \notin N_{2^{k-2}}[b_j]$, then by induction hypothesis applied to position $(G, a_j, b_j, k-1)$, Spoiler wins from position $(G, a_j a_{k-1}, b_j b_{k-1}, k-2)$. Since adding more preselected vertices only helps Spoiler, he would win in particular the remaining game from the positions $(G, \bar{a}a_k a_{k-1}, \bar{b}b_k, \bar{b}_{k-1}, k-2)$. \square

Lemma 3 (see Lemma 3.5 of [GPPT22]). *Consider tuples of vertices $\bar{a}, \bar{a}', \bar{b}, \bar{b}'$ in a graph G such that $\text{dist}(\bar{a}, \bar{a}') > 2^k$ and $\text{dist}(\bar{b}, \bar{b}') > 2^k$. Then $(G, \bar{a}\bar{a}') \cong_k^{\text{local}} (G, \bar{b}\bar{b}')$ if and only if both $(G, \bar{a}) \cong_k^{\text{local}} (G, \bar{b})$ and $(G, \bar{a}') \cong_k^{\text{local}} (G, \bar{b}')$.*

Proof. We prove the statement by induction on k . For $k = 0$, observe that $\text{dist}(\bar{a}, \bar{a}'), \text{dist}(\bar{b}, \bar{b}') > 2^0 = 1$, and thus there are no edges between \bar{a} and \bar{a}' or between \bar{b} and \bar{b}' in G . This means $(\bar{a}\bar{a}', \bar{b}\bar{b}')$ is a partial automorphism if and only if both (\bar{a}, \bar{b}) and (\bar{a}', \bar{b}') are partial automorphisms. This proves the statement for $k = 0$. Next, assume the statement holds for $k-1$ and we will prove it for k .

Assume $(G, \bar{a}) \cong_k^{\text{local}} (G, \bar{b})$ and $(G, \bar{a}') \cong_k^{\text{local}} (G, \bar{b}')$. Note that in particular $(G, \bar{a}') \cong_{k-1}^{\text{local}} (G, \bar{b}')$. We consider the local game at position $(G, \bar{a}\bar{a}', \bar{b}\bar{b}', k)$ and show that Duplicator has a winning strategy. By symmetry, without loss of generality, Spoiler starts with an a -move $a_k \in N_{2^{k-1}}[\bar{a}]$. Duplicator responds according to the winning strategy for the local game at position (G, \bar{a}, \bar{b}, k) yielding $b_k \in N_{2^{k-1}}[\bar{b}]$ such that $(G, \bar{a}a_k) \cong_{k-1}^{\text{local}} (G, \bar{b}b_k)$. By assumption, $\text{dist}(\bar{a}, \bar{a}') > 2^k$ and $\text{dist}(\bar{b}, \bar{b}') > 2^k$, and thus $\text{dist}(\bar{a}a_k, \bar{a}') > 2^{k-1}$ and $\text{dist}(\bar{b}b_k, \bar{b}') > 2^{k-1}$. By induction, since $(G, \bar{a}a_k) \cong_{k-1}^{\text{local}} (G, \bar{b}b_k)$ and $(G, \bar{a}') \cong_{k-1}^{\text{local}} (G, \bar{b}')$, we have $(G, \bar{a}a_k \bar{a}') \cong_{k-1}^{\text{local}} (G, \bar{b}b_k \bar{b}')$. Since we made no assumptions on Spoiler's local move, this implies $(G, \bar{a}\bar{a}') \cong_k^{\text{local}} (G, \bar{b}\bar{b}')$.

Conversely, assume $(G, \bar{a}) \not\cong_k^{\text{local}} (G, \bar{b})$ or $(G, \bar{a}') \not\cong_k^{\text{local}} (G, \bar{b}')$. Without loss of generality, $(G, \bar{a}) \not\cong_k^{\text{local}} (G, \bar{b})$. We consider the local game at position $(G, \bar{a}\bar{a}', \bar{b}\bar{b}', k)$. Spoiler chooses $a_k \in N_{2^{k-1}}[\bar{a}]$ according to his winning strategy at position (G, \bar{a}, \bar{b}, k) . By Lemma 2, Duplicator responds with $b_k \in N_{2^{k-1}}[\bar{b}]$, which is a valid turn in the local game on position (G, \bar{a}, \bar{b}, k) . As Spoiler played according to his winning strategy on that position, we have $(G, \bar{a}a_k) \not\cong_{k-1}^{\text{local}} (G, \bar{b}b_k)$. Again we have $\text{dist}(\bar{a}a_k, \bar{a}') > 2^{k-1}$ and $\text{dist}(\bar{b}b_k, \bar{b}') > 2^{k-1}$ and, by induction, we have $(G, \bar{a}a_k \bar{a}') \not\cong_{k-1}^{\text{local}} (G, \bar{b}b_k \bar{b}')$. Since we made no assumptions on Duplicator's local move this implies $(G, \bar{a}\bar{a}') \not\cong_k^{\text{local}} (G, \bar{b}\bar{b}')$. \square

Theorem 3 (see Lemma 9.4 of [Tor22]). *Consider a graph G with tuples \bar{a}, \bar{b} such that $\text{dist}(\bar{a}, \bar{b}) > 2^{k+1}$. Then*

$$(G, \bar{a}) \cong_k (G, \bar{b}) \iff (G, \bar{a}) \cong_k^{\text{local}} (G, \bar{b}).$$

Proof. The forward direction is easy. Duplicator's winning strategy for the global game when Spoiler makes only local moves is also a winning strategy for the local game, since by Lemma 2 her winning strategy anyways responds locally to local moves.

We prove the backward direction by induction on k . For $k = 0$ the global and local game are the same, hence the statement is true. Assume it holds for $k-1$ and we will prove that it also holds for k . We first prove the following claim.

Claim 1. *If Duplicator has a winning strategy for the game from position (G, \bar{a}, \bar{b}, k) where the first round is local and the remaining rounds are global, then she also has a winning strategy for the global game from position (G, \bar{a}, \bar{b}, k) .*

Proof. We will give a winning strategy for Duplicator for the global game at position (G, \bar{a}, \bar{b}, k) . Without loss of generality, we can assume Spoiler starts the game with an a -move. If Spoiler opens with a local move, then Duplicator can respond according to her given first-local-then-global winning strategy for the position and win the game.

Thus, we may assume that Spoiler opens with a non-local move $a_k \notin N_{2^{k-1}}[\bar{a}]$. We start by arguing that there exists an element $b_k \notin N_{2^{k-1}}[\bar{b}]$ with $(G, a_k) \cong_{k-1}^{\text{local}} (G, b_k)$: if $a_k \notin N_{2^{k-1}}[\bar{b}]$, then we can choose $b_k = a_k$ and are done. Thus assume $a_k \in N_{2^{k-1}}[\bar{b}]$. We will use a role-swapping argument: if Spoiler had played the local element $a_k \in N_{2^{k-1}}[\bar{b}]$ as a b -move (under the name b_k), then Duplicator's winning strategy of the first-local-then-global game would have replied by Lemma 2 with an element $b_k \in N_{2^{k-1}}[\bar{a}]$ as an a -move (under the name a_k). Duplicator would win the remaining $(k-1)$ -round global game, or in other words $(G, \bar{a}a_k) \cong_{k-1} (G, \bar{a}b_k)$. By the forward direction of the theorem, which was already proved above, also $(G, \bar{b}a_k) \cong_{k-1}^{\text{local}} (G, \bar{a}b_k)$. In particular, we also have $(G, a_k) \cong_{k-1}^{\text{local}} (G, b_k)$. Since $b_k \in N_{2^{k-1}}[\bar{a}]$ and $\text{dist}(\bar{a}, \bar{b}) > 2^{k+1} > 2^k$ we have $b_k \notin N_{2^{k-1}}[\bar{b}]$. This yields the desired $b_k \notin N_{2^{k-1}}[\bar{b}]$ with $(G, a_k) \cong_{k-1}^{\text{local}} (G, b_k)$.

Duplicator now responds to Spoiler's move a_k with b_k . Since Duplicator has a k -round winning strategy for the game with preselected tuples \bar{a}, \bar{b} where the first round is local and the remaining rounds are global, she in particular has a $(k-1)$ -round winning strategy for these tuples, that is, $(G, \bar{a}) \cong_{k-1} (G, \bar{b})$, and hence $(G, \bar{a}) \cong_{k-1}^{\text{local}} (G, \bar{b})$ by the forward direction of the theorem. Since $a_k \notin N_{2^{k-1}}[\bar{a}]$, $b_k \notin N_{2^{k-1}}[\bar{b}]$, we have $\text{dist}(\bar{a}, a_k) > 2^{k-1}$ and $\text{dist}(\bar{b}, b_k) > 2^{k-1}$. We can thus apply Lemma 3 for $k-1$ to the tuples $\bar{a}, a_k, \bar{b}, b_k$: Since $(G, a_k) \cong_{k-1}^{\text{local}} (G, b_k)$ and $(G, \bar{a}) \cong_{k-1}^{\text{local}} (G, \bar{b})$, it follows that $(G, \bar{a}a_k) \cong_{k-1}^{\text{local}} (G, \bar{b}b_k)$. By induction hypothesis we have $(G, \bar{a}a_k) \cong_{k-1} (G, \bar{b}b_k)$. Since we made no assumptions on Spoiler's first move, Duplicator has a winning strategy in the global game from position (G, \bar{a}, \bar{b}, k) . \square

We are ready to prove the backwards direction of the statement. Assume $(G, \bar{a}) \cong_k^{\text{local}} (G, \bar{b})$. By the previous claim it suffices to show that Duplicator has a winning strategy for the game from position (G, \bar{a}, \bar{b}, k) where the first round is local and the remaining rounds are global. Hence, let $a_k \in N_{2^{k-1}}[\bar{a}]$ be a local a -move of Spoiler. We let $b_k \in N_{2^{k-1}}[\bar{b}]$ be Duplicator's response that she would play as a winning move in the local game, that is, we have $(G, \bar{a}a_k) \cong_{k-1}^{\text{local}} (G, \bar{b}b_k)$.

Since $\text{dist}(\bar{a}, \bar{b}) > 2^{k+1}$, and $a_k \in N_{2^{k-1}}[\bar{a}]$, $b_k \in N_{2^{k-1}}[\bar{b}]$, it follows that $\text{dist}(\bar{a}a_k, \bar{b}b_k) > 2^k$ (see Figure 2).

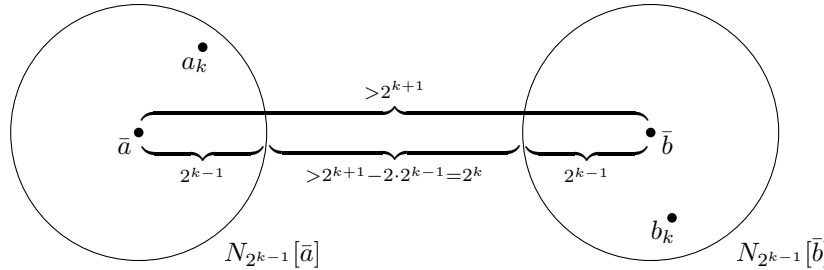


Figure 2: Bounding $\text{dist}(\bar{a}a_k, \bar{b}b_k)$ from below.

By induction hypothesis we have $(G, \bar{a}a_k) \cong_{k-1} (G, \bar{b}b_k)$. As the first move was an arbitrary local a -move, this yields a winning strategy of Duplicator for the game at position (G, \bar{a}, \bar{b}, k) where the first round is local and the remaining rounds are global. By the previous claim, $(G, \bar{a}) \cong_k (G, \bar{b})$. \square

While we do not prove, whether or not the distance requirement $\text{dist}(\bar{a}, \bar{b}) > 2^{k+1}$ in Theorem 3 is tight, the example in Figure 3 illustrates that some form of distance requirement is necessary for the theorem to hold. Let $k = 1$. Then $(G, a) \cong_k^{\text{local}} (G, b)$. However, $(G, a) \not\cong_k (G, b)$ as in the global game, Spoiler can choose the uppermost red element as a global b -move. Duplicator cannot reply with a red element that is not adjacent to a , and hence loses the game.

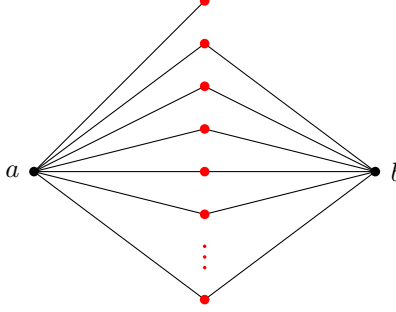


Figure 3: An example illustrating the need for a distance constraint in Theorem 3.

4.5 Local Games and Local Types

We now establish the connection between local games and local first-order logic. Unlike in Gaifman's Locality Theorem, we do not increase the quantifier rank when localizing formulas.

Let G be a graph and \bar{a} be a tuple of vertices of G . It is well known that $\text{tp}_q(G, \bar{a}) = \text{tp}_q(G, \bar{b})$ if and only if $(G, \bar{a}) \cong_q (G, \bar{b})$. The *localization* of a formula φ with free variables is the formula with the same free variables as φ that replaces every subformula $\exists x \psi(x, \bar{y})$ with $\exists x \in N_{2^{k-1}}[\bar{y}] \psi(x, \bar{y})$ (or more precisely $\exists x x \in N_{2^{k-1}}[\bar{y}] \wedge \psi(x, \bar{y})$). Likewise, every subformula $\forall x \psi(x, \bar{y})$ is replaced with $\forall x \in N_{2^{k-1}}[\bar{y}] \psi(x, \bar{y})$ (or more precisely $\forall x x \in N_{2^{k-1}}[\bar{y}] \rightarrow \psi(x, \bar{y})$).

We call a formula *local* if it is the localization of some formula. As shown in the following lemma, one can express with $k-1$ quantifiers that $x \in N_{2^{k-1}}[\bar{y}]$, and thus localizing a formula does not change its quantifier rank.

Lemma 4. *There exists a formula with quantifier rank k and free variables $x\bar{y}$ expressing that $x \in N_{2^k}[\bar{y}]$.*

Proof. We can check whether $x \in N_{2^0}[\bar{y}]$ using the quantifier-free formula $\bigvee_{y \in \bar{y}} E(x, y) \vee x = y$. For $k > 0$, we note that $x \in N_{2^k}[\bar{y}]$ if and only if $\exists z (z \in N_{2^{k-1}}[x] \wedge z \in N_{2^{k-1}}[\bar{y}])$. \square

Let G be a graph and $\bar{a} \in V(G)^{|\bar{a}|}$ be a tuple in G . We partition the finite set of all (normalized) local formulas $\varphi(\bar{x})$ with $|\bar{x}| = |\bar{a}|$ and quantifier rank at most q over the signature of G into the sets $\text{tp}_q^{\text{local}}(G, \bar{a})$ and $\overline{\text{tp}}_q^{\text{local}}(G, \bar{a})$ such that $\varphi(\bar{x}) \in \text{tp}_q^{\text{local}}(G, \bar{a})$ if and only if $G \models \varphi(\bar{a})$ and conversely $\varphi(\bar{x}) \in \overline{\text{tp}}_q^{\text{local}}(G, \bar{a})$ if and only if $G \not\models \varphi(\bar{a})$. We call $\text{tp}_q^{\text{local}}(G, \bar{a})$ the *local q -type of \bar{a} in G* .

We next relate local types and local games.

Lemma 5. *If $\text{tp}_k^{\text{local}}(G, \bar{a}) = \text{tp}_k^{\text{local}}(G, \bar{b})$, then $(G, \bar{a}) \cong_k^{\text{local}} (G, \bar{b})$.*

Proof. We prove the claim by induction on k . For $k = 0$, $\text{tp}_0^{\text{local}}(G, \bar{a})$ and $\text{tp}_0^{\text{local}}(G, \bar{b})$ are known as the atomic types of \bar{a} and \bar{b} . These are equal if and only if the mapping $\bar{a} \mapsto \bar{b}$ is a partial isomorphism, which in turn is equivalent to $(G, \bar{a}) \cong_0^{\text{local}} (G, \bar{b})$.

Let us assume that the statement holds for $k - 1$ and show that it also holds for k . Consider the local game at position (G, \bar{a}, \bar{b}, k) . Without loss of generality, Spoiler starts the game by an a -move $a_k \in N_{2k-1}(\bar{a})$. Let

$$\tau(\bar{y}, x) = \bigwedge_{\varphi(\bar{y}x) \in \text{tp}_{k-1}^{\text{local}}(G, \bar{a}a_k)} \varphi(\bar{y}x) \wedge \bigwedge_{\varphi(\bar{y}x) \in \overline{\text{tp}}_{k-1}^{\text{local}}(G, \bar{a}a_k)} \neg \varphi(\bar{y}x)$$

be the local formula that exactly captures $\text{tp}_{k-1}^{\text{local}}(G, \bar{a}a_k)$. More precisely, for every tuple $\bar{a}'a'_k \in V(G)^{|a|+1}$ we have $\text{tp}_{k-1}^{\text{local}}(G, \bar{a}'a'_k) = \text{tp}_{k-1}^{\text{local}}(G, \bar{a}a_k)$ if and only if $G \models \tau(\bar{a}', a'_k)$.

The formula $\psi(\bar{y}) = \exists x \in N_{2k-1}[\bar{y}] \tau(\bar{y}x)$ is a local formula with quantifier rank k , and is therefore contained in $\text{tp}_k^{\text{local}}(G, \bar{a})$ as witnessed by instantiating x with a_k . By assumption on equality of local k -types we then also have that $\psi(\bar{y}) \in \text{tp}_k^{\text{local}}(G, \bar{b})$ and hence there exists an element $b_k \in N_{2k-1}(\bar{b})$ such that $\text{tp}_{k-1}^{\text{local}}(G, \bar{b}b_k) = \text{tp}_{k-1}^{\text{local}}(G, \bar{a}a_k)$. Duplicator chooses b_k as her response. The remaining game continues from position $(G, \bar{a}a_k, \bar{b}b_k, k - 1)$. Since $\text{tp}_{k-1}^{\text{local}}(G, \bar{b}b_k) = \text{tp}_{k-1}^{\text{local}}(G, \bar{a}a_k)$ Duplicator wins by induction hypothesis. \square

If is not difficult to prove that also the converse of the lemma is true, however, we refrain from giving the proof as it is not needed for our further argumentation.

4.6 Games and Types with Guards

Theorem 3 and Lemma 5 together already show that for tuples of sufficiently large distance equality of local types implies equality of global types. We will need a stronger statement for graphs where a specific set of vertices is highlighted. To this end, we introduce special starting positions (G, A, B, k) , where $A, B \subseteq V(G)$ are sets of vertices (for the global and local EF-game), which we call *guards*. Spoiler and Duplicator select elements a_k and b_k in the usual way with the constraints $a_k \in A$ and $b_k \in B$, and afterwards the (global or local) game continues at position $(G, a_k, b_k, k - 1)$ as usual. Hence, both for the global and the local game, the role of the sets A and B is merely to constrain (guard) the choices for the first round. We write $(G, A) \cong_k (G, B)$ or $(G, A) \cong_k^{\text{local}} (G, B)$ if Duplicator has a winning strategy for the global or local game starting from position (G, A, B, k) .

First, we extend Lemma 5 to our new starting positions. Note that the following theorem no longer mentions local types, but global types of neighborhoods. Recall that $G\langle X \mapsto W \rangle$ denotes the graph G with the additional color predicate X interpreted as the vertex set $W \subseteq V(G)$.

Lemma 6. *If $\text{tp}_k(G\langle X \mapsto A \rangle [N_{2k-1-1}[A]]) = \text{tp}_k(G\langle X \mapsto B \rangle [N_{2k-1-1}[B]])$, then $(G, A) \cong_k^{\text{local}} (G, B)$.*

Proof. Fix G, A and B with $\text{tp}_k(G\langle X \mapsto A \rangle [N_{2k-1-1}[A]]) = \text{tp}_k(G\langle X \mapsto B \rangle [N_{2k-1-1}[B]])$. For brevity, let $G_A := G\langle X \mapsto A \rangle$ and $G_B := G\langle X \mapsto B \rangle$. To prove the statement, we need the following observation about local formulas.

Claim 2. *Let $\varphi(x)$ be a local formula with quantifier rank at most $k - 1$. Then $G_A \models \exists x \in X \varphi(x)$ if and only if $G_B \models \exists x \in X \varphi(x)$.*

Proof. Since $G_A [N_{2k-1-1}[A]]$ and $G_B [N_{2k-1-1}[B]]$ have the same k -type, they agree in their evaluation of the sentence $\exists x \in X \varphi(x)$ with quantifier rank at most k .

Since $\varphi(\bar{x})$ is local and has quantifier rank $k - 1$, all the quantified variables in $\varphi(x)$ can only lie within distance at most $\sum_{i=1}^{k-1} 2^{i-1} = \sum_{i=0}^{k-2} 2^i = 2^{k-1} - 1$ from x . Hence, all variables quantified in $\exists x \in X \varphi(x)$ must lie within distance at most $2^{k-1} - 1$ from X . Therefore, evaluating it on $G_A [N_{2k-1-1}[A]]$ and G_A yields the same answer. The same holds for $G_B [N_{2k-1-1}[B]]$ and G_B . Hence, also G_A and G_B agree in their evaluation of $\exists x \in X \varphi(x)$. \square

Consider the local game at position (G, A, B, k) . Without loss of generality, Spoiler starts the game by an a -move $a_k \in A$. Let

$$\tau(x) = \bigwedge_{\varphi(x) \in \text{tp}_{k-1}^{\text{local}}(G_A, a_k)} \varphi(x) \wedge \bigwedge_{\varphi(x) \in \overline{\text{tp}}_{k-1}^{\text{local}}(G_A, a_k)} \neg \varphi(x)$$

be the local formula that defines $\text{tp}_{k-1}^{\text{local}}(G_A, a_k)$.

The sentence $\exists x \in X \tau(x)$ holds on G_A , as witnessed by instantiating x with a_k . By Claim 2, as $\tau(x)$ has quantifier rank $k-1$, $\exists x \in X \tau(x)$ is also true on G_B . Hence, there exists an element $b_k \in B$ such that $\text{tp}_{k-1}^{\text{local}}(G_A, a_k) = \text{tp}_{k-1}^{\text{local}}(G_B, b_k)$. Then in particular $\text{tp}_{k-1}^{\text{local}}(G, a_k) = \text{tp}_{k-1}^{\text{local}}(G, b_k)$. Duplicator chooses b_k as his next element. Then by Lemma 5 we have $(G, a_k) \cong_{k-1}^{\text{local}} (G, b_k)$. Since we made no assumptions on Spoiler's first move $a_k \in A$, and Duplicator's response always yields a $b_k \in B$, we now have $(G, A) \cong_k^{\text{local}} (G, B)$ as desired. \square

Since for starting positions (G, A, B, k) , the local and global game allow the same first moves, we get the following simple consequence of Theorem 3.

Lemma 7. *Consider a graph G with sets $A, B \subseteq V(G)$ such that $\text{dist}(A, B) > 2^k$. Then $(G, A) \cong_k (G, B)$ if and only if $(G, A) \cong_k^{\text{local}} (G, B)$.*

We can use these new starting positions to determine the truth values of formulas in graphs where the starting sets are highlighted.

Lemma 8. *Assume $(G, A) \cong_k (G, B)$. Then for every formula $\varphi(x)$ of quantifier rank at most $k-1$ in the signature of G we have $G\langle A \rangle \models \exists x \in A \varphi(x) \iff G\langle B \rangle \models \exists x \in B \varphi(x)$.*

Proof. Assume $G\langle A \rangle \models \exists x \in A \varphi(x)$, that is, there exists $a_k \in A$ with $G \models \varphi(a_k)$. Spoiler chooses $a_k \in A$ and Duplicator responds with $b_k \in B$ such that $(G, a_k) \cong_{k-1} (G, b_k)$. Hence, $\text{tp}_{k-1}(G, a_k) = \text{tp}_{k-1}(G, b_k)$, and in particular $G \models \varphi(b_k)$. We have $G\langle B \rangle \models \exists x \in B \varphi(x)$. The converse holds by symmetry. \square

We combine Lemma 6, Lemma 7 and Lemma 8 into the following statement.

Lemma 9. *Assume $\text{dist}(A, B) > 2^k$ and*

$$\text{tp}_k(G\langle X \mapsto A \rangle [N_{2^{k-1}-1}[A]]) = \text{tp}_k(G\langle X \mapsto B \rangle [N_{2^{k-1}-1}[B]]).$$

Then for every formula $\varphi(x)$ of quantifier rank at most $k-1$ in the signature of G we have $G\langle A \rangle \models \exists x \in A \varphi(x) \iff G\langle B \rangle \models \exists x \in B \varphi(x)$.

Again, Figure 3 illustrates that the distance constraint is necessary. For $k=2$, $A = \{a\}$ and $B = \{b\}$, we have that both $G\langle X \mapsto A \rangle [N_{2^{2-1}-1}[A]]$ and $G\langle X \mapsto B \rangle [N_{2^{2-1}-1}[B]]$ have the same type: both are a star, whose center is marked X and whose leaves are marked red. However, for $\varphi(x) := \forall y \text{ Red}(y) \rightarrow E(x, y)$ we have

$$G\langle A \rangle \models \exists x \in A \varphi(x) \text{ and } G\langle B \rangle \not\models \exists x \in B \varphi(x).$$

Extending the statement to accommodate for further free variables in φ , will yield Theorem 4, which we restate for convenience.

Theorem 4. *Let G be a graph and let $A, B \subseteq V(G)$ be vertex sets such that $\text{dist}(A, B) > 2^k$ and $\text{tp}_k(G\langle X \mapsto A \rangle [N_{2^{k-1}-1}[A]]) = \text{tp}_k(G\langle X \mapsto B \rangle [N_{2^{k-1}-1}[B]])$. Let $\bar{w} \in V(G)^{|\bar{y}|}$ be vertices with $\text{dist}(\bar{w}, A \cup B) \geq 2^k$. Then for every formula $\varphi(\bar{y}, x)$ of quantifier rank at most $k-1$ in the signature of G we have $G\langle A \rangle \models \exists x \in A \varphi(\bar{w}, x) \iff G\langle B \rangle \models \exists x \in B \varphi(\bar{w}, x)$.*

Proof. Assume $\bar{w} = w_1, \dots, w_\ell$. We define G' to be the graph extended with 2ℓ new color predicates W_i and N_i for $1 \leq i \leq \ell$. We interpret $W_i = \{w_i\}$ and $N_i = N[w_i] \setminus \{w_i\}$ for $1 \leq i \leq \ell$. We define $\varphi'(x)$ to be the formula obtained from $\varphi(\bar{y}, x)$ by replacing all atoms $E(w_i, z)$ and $E(z, w_i)$ with $N_i(z)$ and all atoms $(w_i = z)$ and $(z = w_i)$ with $W_i(z)$. Then for every $v \in V(G)$ we have $G \models \varphi(\bar{w}, v) \iff G' \models \varphi'(v)$ and thus it is sufficient to show $G' \langle A \rangle \models \exists x \in A \varphi'(x) \iff G' \langle B \rangle \models \exists x \in B \varphi'(x)$. Since $\text{dist}(\bar{w}, A) \geq 2^k$, we have $G[N_{2^{k-1}-1}[A]] = G'[N_{2^{k-1}-1}[A]]$. The same holds for B and thus $\text{tp}_k(G'[N_{2^{k-1}-1}(A)] \langle A \mapsto X \rangle) = \text{tp}_k(G'[N_{2^{k-1}-1}(B)] \langle B \mapsto X \rangle)$. The statement then follows from Lemma 9. \square

5 Model Checking

In this section, we present our model checking theorem for structurally nowhere dense graph classes.

Theorem 1. *For every structurally nowhere dense class \mathcal{C} there exists a function $f: \mathbb{N} \rightarrow \mathbb{N}$ such that, given a graph $G \in \mathcal{C}$ and sentence φ , one can decide whether $G \models \varphi$ in time $f(|\varphi|) \cdot |V(G)|^{11}$.*

As a stepping stone, we prove the following conditional theorem for monadically stable classes of graphs.

Theorem 5. *Let \mathcal{C} be a monadically stable graph class admitting flip-closed sparse weak neighborhood covers. There exists a function $f: \mathbb{N} \rightarrow \mathbb{N}$ such that, given a graph $G \in \mathcal{C}$ and sentence φ , one can decide whether $G \models \varphi$ in time $f(|\varphi|) \cdot |V(G)|^{11}$.*

Note that the condition on \mathcal{C} is an existential statement: if \mathcal{C} admits flip-closed sparse weak neighborhood covers, then we can solve the model checking problem on \mathcal{C} efficiently. To actually calculate the required covers during our algorithm, we will make use of the following theorem whose proof is deferred to Section 6.

Theorem 10. *There is an algorithm that gets as input an n -vertex graph G and numbers $r, s \in \mathbb{N}$ and computes in time $\mathcal{O}(n^{9.8})$ a weak r -neighborhood cover with degree $\mathcal{O}(\log(n)^2 + 1)d^*$ and spread s , where d^* is the smallest number such that G admits a weak r -neighborhood cover with degree d^* and spread s .*

For structurally nowhere dense classes, we are able to prove the existence of the desired covers as stated in the following theorem, which we will prove in Section 7.

Theorem 11. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon)$ such that for every $G \in \mathcal{C}$ there exists a weak r -neighborhood cover with degree at most $c(r, \varepsilon) \cdot |G|^\varepsilon$ and spread at most $34r$. In particular, \mathcal{C} admits flip-closed sparse weak neighborhood covers.*

As every structurally nowhere dense class is monadically stable, combining Theorem 5 and Theorem 11 now yields Theorem 1.

5.1 Setup

Recall that $\text{game-depth}(\mathcal{C}, \rho)$ is the smallest number such that the Flipper strategy flip^* wins the radius- ρ Flipper game on \mathcal{C} in $\text{game-depth}(\mathcal{C}, \rho)$ rounds. This means, graphs resulting from $\text{game-depth}(\mathcal{C}, \rho)$ many rounds of play by flip^* are winning positions for Flipper, that is, single vertices. Here, model checking is trivial. We assume an algorithm for graphs resulting from $\ell + 1$

rounds of play (with the precise definition given by the following Definition 4), and use it to also do model checking on for graphs with only ℓ rounds played. Repeating this procedure gives us an algorithm for graphs on which zero rounds have been played, that is, a model checking algorithm for all graphs from \mathcal{C} . The choice $\rho := (6q\sigma(2^q) + 1)(2^q + 1)$ for the radius of the game emerges from the details of our proofs.

Definition 4. Let \mathcal{C} be a monadically stable graph class admitting flip-closed sparse weak neighborhood covers with spread σ , and let $q, \ell, c \in \mathbb{N}$. We choose a radius $\rho := (6q\sigma(2^q) + 1)(2^q + 1)$ for the Flipper game. Note that ρ depends only on \mathcal{C} and q . Consider an algorithm that gets as input

- a (\mathcal{C}, ρ) -history $(G_0, \mathcal{I}_0), \dots, (G_\ell, \mathcal{I}_\ell)$ of length ℓ from the Flipper game,
- a coloring G of G_ℓ with a signature of at most c colors, and
- a sentence φ with quantifier rank at most q

and decides whether $G \models \varphi$. We say this is an *efficient* $\text{MC}(\mathcal{C}, q, \ell, c)$ -algorithm, if there exists a function f_{MC} bounding the runtime for every $\varepsilon > 0$ by

$$f_{\text{MC}}(q, \ell, c, \varepsilon) \cdot |V(G)|^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8},$$

where $d := \text{game-depth}(\mathcal{C}, \rho) - \ell$ bounds the number of rounds needed to win the remaining Flipper game.

5.2 Computing Guarded Formulas

As central building block of our algorithm, the following theorem converts sentences into guarded sentences, assuming we already have an efficient model checking algorithm for graphs where the game has progressed by one extra round.

Theorem 6. Let \mathcal{C} be a monadically stable graph class admitting flip-closed sparse weak neighborhood covers with spread σ , and let $\rho = (6q\sigma(2^q) + 1)(2^q + 1)$. Given as input

- a (\mathcal{C}, ρ) -history $\mathcal{H} = (G_0, \mathcal{I}_0), \dots, (G_\ell, \mathcal{I}_\ell)$ of length ℓ ,
- a coloring G of G_ℓ with a signature of at most c colors,
- a sentence φ with quantifier rank at most q , and
- an efficient $\text{MC}(\mathcal{C}, q, \ell + 1, c + 3)$ -algorithm,

one can compute sets $U_1, \dots, U_t \subseteq V(G)$, for some constant t depending only on q and c , as well as a (U_1, \dots, U_t) -guarded sentence ξ of quantifier rank q . Each U_i is contained in an $(q \cdot 3\sigma(2^q))$ -neighborhood of G and

$$G \models \varphi \iff G\langle U_1, \dots, U_t \rangle \models \xi.$$

There exists a function $f(q, c, \ell, \varepsilon)$ such that for every $\varepsilon > 0$, the running time of this procedure is bounded by

$$f(q, c, \ell, \varepsilon) \cdot |V(G)|^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8},$$

where $d := \text{game-depth}(\mathcal{C}, \rho) - \ell$ bounds the number of rounds needed to win the remaining Flipper game.

Instead of guarding all quantifiers at once, we start with guarding only one outermost quantifier. The following theorem will be the central step of our construction.

Theorem 7. Let \mathcal{C} be a monadically stable graph class admitting flip-closed sparse weak neighborhood covers with spread σ , and let $\rho = (6q\sigma(2^q) + 1)(2^q + 1)$. Given as input

- a (\mathcal{C}, ρ) -history $\mathcal{H} = (G_0, \mathcal{I}_0), \dots, (G_\ell, \mathcal{I}_\ell)$ of length ℓ ,
- a coloring G of G_ℓ with a signature of at most c colors,
- a formula $\exists x \varphi(\bar{y}, x)$ of quantifier rank at most q ,
- sets $W_1, \dots, W_{|\bar{y}|}$, each contained in an r -neighborhood of G , and
- an efficient $\text{MC}(\mathcal{C}, q, \ell + 1, c + 3)$ -algorithm,

one can compute sets $U_1, \dots, U_t \subseteq V(G)$, for some constant t depending only on q , c , and $|\bar{y}|$. Each U_i is contained in an $(r + 3\sigma(2^q))$ -neighborhood of G and for all tuples $\bar{w} \in W_1 \times \dots \times W_{|\bar{y}|}$, we have

$$G \models \exists x \varphi(\bar{w}, x) \iff \bigvee_{i=1}^t G[U_i] \models \exists x \in U_i \varphi(\bar{w}, x).$$

There exists a function $f(q, c, \ell, \varepsilon, |\bar{y}|)$ such that for every $\varepsilon > 0$, the running time of this procedure is bounded by

$$f(q, c, \ell, \varepsilon, |\bar{y}|) \cdot |V(G)|^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8},$$

where $d := \text{game-depth}(\mathcal{C}, \rho) - \ell$ bounds the number of rounds needed to win the remaining Flipper game.

Proof. Let n be the number of vertices of G . Our goal is to compute the set of guards $\mathcal{U} = \{U_1, \dots, U_t\}$.

Neighborhood Cover Computation. We use Theorem 10 to compute in time $\mathcal{O}(n^{9.8})$ a weak 2^q -neighborhood cover of G with degree $\mathcal{O}(\log(n)^2 + 1) \cdot d^*$ and spread $\sigma(2^q)$, where d^* is the smallest number such that G admits a weak 2^q -neighborhood cover with degree d^* and spread $\sigma(2^q)$. Let us argue the existence of a function $g_1(q, \varepsilon, \ell)$ such that for every $\varepsilon > 0$, this computes a cover of degree $g_1(q, \varepsilon, \ell) \cdot n^\varepsilon$.

Let $\varepsilon > 0$. The graph G is obtained from $G_0 \in \mathcal{C}$ by performing at most ℓ flips and removing vertices. Hence, by Definition 3, there exists a function $g(q, \varepsilon, \ell)$ such that G has a weak 2^q -neighborhood cover with degree $g(q, \varepsilon, \ell) \cdot n^{\varepsilon/2}$ and spread $\sigma(2^q)$. Hence, $d^* \leq g(q, \varepsilon, \ell) \cdot n^{\varepsilon/2}$ and the computed neighborhood cover has degree at most $\mathcal{O}(\log(n)^2 + 1) \cdot g(q, \varepsilon, \ell) \cdot n^{\varepsilon/2}$. Since logarithmic factors are dominated by any polynomial factor, this can be bounded by $g_1(q, \varepsilon, \ell) \cdot n^\varepsilon$ for some appropriately chosen function $g_1(q, \varepsilon, \ell)$.

Let $\{C_1, \dots, C_m\}$ be the computed weak 2^q -neighborhood cover. Without loss of generality, we can assume $m \leq n$, since otherwise redundant sets can be removed. We partition the vertices of G into sets V_1, \dots, V_m such that for all $v \in V_i$, $N_{2^q}[v] \subseteq C_i$. Ties are broken arbitrarily.

Splitting the Existential Quantifier. It will be useful to partition the existential quantification of x in our input formula $\exists x \varphi(\bar{y}, x)$ into a quantification over sets that are near and that are far from $W_1, \dots, W_{|\bar{y}|}$. To this end, let $V'_i := V_i \setminus N_{2^q} \left[\bigcup_{k=1}^{|\bar{y}|} W_k \right]$. Since every vertex of G is in some V_i ,

for all tuples $\bar{w} \in W_1 \times \dots \times W_{|y|}$

$$G \models \exists x \varphi(\bar{w}, x) \iff \bigvee_{i=1}^{|\bar{y}|} G\langle N_{2^q}[W_i] \rangle \models \exists x \in N_{2^q}[W_i] \varphi(\bar{w}, x) \vee \bigvee_{i=1}^m G\langle V'_i \rangle \models \exists x \in V'_i \varphi(\bar{w}, x). \quad (1)$$

Remember that the size of our solution \mathcal{U} may depend only on q , c , and $|\bar{y}|$. Adding the sets $N_{2^q}[W_1], \dots, N_{2^q}[W_{|\bar{y}|}]$ to \mathcal{U} would respect this size constraint. However, since m may depend on n , we are not allowed to add all sets V'_1, \dots, V'_m to \mathcal{U} . In the remainder of this proof, we will use Theorem 4 and the fact that each V'_i is sufficiently far away from $W_1, \dots, W_{|\bar{y}|}$ to construct a set $S \subseteq [m]$ with the following property.

Property 1. *The size of $S \subseteq [m]$ depends only on q and c and for all tuples $\bar{w} \in W_1 \times \dots \times W_{|\bar{y}|}$*

$$\bigvee_{i=1}^m G\langle V'_i \rangle \models \exists x \in V'_i \varphi(\bar{w}, x) \implies \bigvee_{i \in S} G\langle X \mapsto N_{2^q+2\sigma(2^q)}[V'_i] \rangle \models \exists x \in X \varphi(\bar{w}, x).$$

After we found such a set S , we set

$$\mathcal{U} = \{N_{2^q}[W_1], \dots, N_{2^q}[W_{|\bar{y}|}]\} \cup \{N_{2^q+2\sigma(2^q)}[V'_i] \mid i \in S\}.$$

Note that $|\mathcal{U}|$ depends only on q , c , and $|\bar{y}|$. Combining (1) and Property 1, it holds for all tuples $\bar{w} \in W_1 \times \dots \times W_{|\bar{y}|}$ that

$$G \models \exists x \varphi(\bar{w}, x) \implies \bigvee_{U \in \mathcal{U}} G\langle U \rangle \models \exists x \in U \varphi(\bar{w}, x).$$

The backwards implication of this statement holds obviously, since the right-hand side merely restricts the quantification of x . This yields the central statement

$$G \models \exists x \varphi(\bar{w}, x) \iff \bigvee_{U \in \mathcal{U}} G\langle U \rangle \models \exists x \in U \varphi(\bar{w}, x).$$

Since each W_i is contained in an r -neighborhood of G , each $N_{2^q}[W_i]$ is contained in an $(r + 2^q)$ -neighborhood and (with $\sigma(2^q) \geq 2^q$) also in an $(r + \sigma(2^q))$ -neighborhood. Each set $N_{2^q}[V'_i]$ is contained in C_i , which by construction is contained in an $\sigma(2^q)$ -neighborhood of G . It follows that $N_{2^q+2\sigma(2^q)}[V'_i] = N_{2\sigma(2^q)}[N_{2^q}[V'_i]]$ is contained in a $3\sigma(2^q)$ -neighborhood in G . Hence, each $U \in \mathcal{U}$ is contained in an $(r + 3\sigma(2^q))$ -neighborhood of G . To finish the proof, we have to compute a small representative set S with Property 1.

Flip and Type Computation. As a first step towards computing S , we show how to use our given efficient $\text{MC}(\mathcal{C}, q, \ell + 1, c + 3)$ -algorithm to compute $\text{tp}_q(G[N_{2^{q-1}-1}[V'_i]] \langle X \mapsto V'_i \rangle)$ for all $i \in [m]$. To this end, we do for every $i \in [m]$ the following computations. Let $H_{\ell+1} := G[N_{2^{q-1}-1}[V'_i]]$. Note that this corresponds to a Connector move in the radius- $\sigma(2^q) \leq \rho$ Flipper game. We apply the radius- ρ Flipper strategy flip^* (for the class $\mathcal{C} \ni G_0$) to the graph $H_{\ell+1}$ and internal state \mathcal{I}_ℓ , yielding a flip F and a new internal state $\mathcal{I}_{\ell+1}$. By Theorem 2, this takes time

$$g_2(q) \cdot |V(G_0)|^2,$$

for some function $g_2(q)$. Let $G_{\ell+1} := H_{\ell+1} \oplus F$. We can now extend \mathcal{H} to a (\mathcal{C}, ρ) -history of length $\ell + 1$ by appending the new pair $(G_{\ell+1}, \mathcal{I}_{\ell+1})$. We spend 3 additional colors to construct $G_{\ell+1}^+$ by marking in $G_{\ell+1}$ with unary predicates the two flip sets from F , as well as the vertices from V'_i . Next, we enumerate the set Φ of normalized first-order sentences with quantifier rank at most q over the signature of $G_{\ell+1}^+$. Recall that $|\Phi|$ is bounded by a function of q and c . We use the given efficient $\text{MC}(\mathcal{C}, q, \ell + 1, c + 3)$ -algorithm to evaluate every formula from Φ on $G_{\ell+1}^+$ and therefore compute $\text{tp}_q(G_{\ell+1}^+)$ in time

$$g_3(q, c) \cdot f_{\text{MC}}(q, \ell + 1, c + 3, \varepsilon) \cdot |V(G_{\ell+1}^+)|^{((1+\varepsilon)^{d-1})} \cdot |V(G_0)|^{9.8},$$

for some function $g_3(q, c)$. Let us now argue how to derive $\text{tp}_q(G[N_{2^{q-1}-1}[V'_i]]\langle V'_i \rangle)$ from $\text{tp}_q(G_{\ell+1}^+)$. This is easy to do by observing that for every sentence ψ , we have $\psi \in \text{tp}_q(G[N_{2^{q-1}-1}[V'_i]]\langle V'_i \rangle)$ if and only if $\psi' \in \text{tp}_q(G_{\ell+1}^+)$ where ψ' is obtained from ψ by substituting every occurrence of the edge relation $E(x, y)$ with $E(x, y) \oplus ((x \in A \wedge y \in B) \vee (x \in B \wedge y \in A))$ where A and B are the color predicates marking the flip sets of F . Similarly, we can derive $\text{tp}_q(G[N_{2^{q-1}-1}[V'_i]]\langle X \mapsto V'_i \rangle)$.

Computing a Representative Set. Now we use the previously computed q -types to pick S as a minimal subset of $[m]$ such that

$$\{\text{tp}_q(G[N_{2^{q-1}-1}[V'_i]]\langle X \mapsto V'_i \rangle) \mid i \in [m]\} = \{\text{tp}_q(G[N_{2^{q-1}-1}[V'_i]]\langle X \mapsto V'_i \rangle) \mid i \in S\}.$$

The size of S is at most the number of possible q -types on graphs with $c + 3$ colors, and thus can be bounded as a function of q and c . In order to show that S satisfies Property 1, let us fix $\bar{w} \in W_1 \times \dots \times W_{|\bar{y}|}$ and argue that

$$\bigvee_{i=1}^m G\langle V'_i \rangle \models \exists x \in V'_i \varphi(\bar{w}, x) \implies \bigvee_{i \in S} G\langle X \mapsto N_{2^q+2\sigma(2^q)}(V'_i) \rangle \models \exists x \in X \varphi(\bar{w}, x).$$

Assume $G\langle V'_i \rangle \models \exists x \in V'_i \varphi(\bar{w}, x)$ for some i . If $V'_i \subseteq \bigcup_{j \in S} N_{2^q+2\sigma(2^q)}[V'_j]$ for some $j \in S$, then the right-hand side follows immediately, so we can assume $V'_i \not\subseteq \bigcup_{j \in S} N_{2^q+2\sigma(2^q)}[V'_j]$ for all $j \in S$. Fix some $j \in S$ and let us show that $\text{dist}(V'_i, V'_j) > 2^q$ and $\text{dist}(\bar{w}, V'_i \cup V'_j) > 2^q$. Since we have $V'_i \not\subseteq \bigcup_{j \in S} N_{2^q+2\sigma(2^q)}[V'_j]$, there exists a vertex in V'_i that has distance greater than $2^q + 2\sigma(2^q)$ from every vertex in V'_j . Since V'_i embeds in a subgraph of G with diameter at most $2\sigma(2^q)$, every vertex in V'_i has distance greater than 2^q from every vertex in V'_j . This means $\text{dist}(V'_i, V'_j) > 2^q$. We finally establish $\text{dist}(\bar{w}, V'_i \cup V'_j) > 2^q$ by combining

$$V'_i := V_i \setminus N_{2^q} \left[\bigcup_{k=1}^{|\bar{y}|} W_k \right], \quad V'_j := V_j \setminus N_{2^q} \left[\bigcup_{k=1}^{|\bar{y}|} W_k \right], \quad \bar{w} \in W_1 \times \dots \times W_{|\bar{y}|}.$$

The set S was chosen representative in the sense that there is some $j \in S$ with

$$\text{tp}_q(G[N_{2^{q-1}-1}[V'_i]]\langle V'_i \rightarrow X \rangle) = \text{tp}_q(G[N_{2^{q-1}-1}[V'_j]]\langle V'_j \rightarrow X \rangle).$$

Since $G\langle V'_i \rangle \models \exists x \in V'_i \varphi(\bar{w}, x)$, by Theorem 4, also $G\langle V'_j \rangle \models \exists x \in V'_j \varphi(\bar{w}, x)$ and the right-hand side holds. Hence, S satisfies Property 1.

Running Time Analysis. At first, we analyze the running time spent for the computations in the paragraph *Flip and Type Computation*. As stated there, the run time is (using $m \leq n$) bounded by

$$\sum_{i \in [m]} g_2(q) \cdot |V(G_0)|^2 \leq n \cdot g_2(q) \cdot |V(G_0)|^2 \quad (2)$$

for computing the flips, plus

$$\sum_{i \in [m]} g_3(q, c) \cdot f_{MC}(q, \ell + 1, c + 3, \varepsilon) \cdot |N_{2^q}[V'_i]|^{((1+\varepsilon)^{d-1})} \cdot |V(G_0)|^{9.8}$$

for computing the q -types. Note that for all $\alpha \geq 1$ and non-negative numbers n_1, \dots, n_m we have $\sum_{i \in [m]} n_i^\alpha \leq (\sum_{i \in [m]} n_i)^\alpha$, bounding the running time for the q -type computation by

$$g_3(q, c) \cdot f_{MC}(q, \ell + 1, c + 3, \varepsilon) \cdot \left(\sum_{i \in [m]} |N_{2^q}[V'_i]| \right)^{((1+\varepsilon)^{d-1})} \cdot |V(G_0)|^{9.8}.$$

For every $i \in m$, we have $N_{2^q}[V'_i] \subseteq N_{2^q}[V_i] \subseteq C_i$, yielding

$$\sum_{i \in [m]} |N_{2^q}[V'_i]| \leq \sum_{i \in [m]} |C_i| \leq g_1(q, \varepsilon, \ell) \cdot n^{1+\varepsilon},$$

where the last bound follows from the fact that we have n vertices, each occurring in at most $g_1(q, \varepsilon, \ell) \cdot n^\varepsilon$ clusters of the cover $\{C_1, \dots, C_m\}$. Combining the previous two inequalities bounds the running time of the type computation by

$$g_3(q, c) \cdot f_{MC}(q, \ell + 1, c + 3, \varepsilon) \cdot \left(g_1(q, \varepsilon, \ell) \cdot n^{1+\varepsilon} \right)^{((1+\varepsilon)^{d-1})} \cdot |V(G_0)|^{9.8},$$

which is equal to

$$g_3(q, c) \cdot f_{MC}(q, \ell + 1, c + 3, \varepsilon) \cdot g_1(q, \varepsilon, \ell)^{((1+\varepsilon)^{d-1})} \cdot n^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8}. \quad (3)$$

The total running time spent in this paragraph, as given by the sum of (2) and (3) can be bounded by $g_4(q, \varepsilon, \ell) \cdot n^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8}$, for some function $g_4(q, \varepsilon, \ell)$.

The computation in the paragraph *Neighborhood Cover Computation* takes time $\mathcal{O}(n^{9.8})$. Since the size of the representative set is bounded by a function of q and c , we can bound the computation time for the paragraphs *Splitting the Existential Quantifier* and *Computing a Representative Set* by $g_5(q, c, |\bar{y}|) \cdot n^2$, for some function $g_5(q, c, |\bar{y}|)$. Since $n \leq |V(G_0)|$, we can choose a function $f(q, c, \ell, \varepsilon, |\bar{y}|)$ such that the total running time is bounded by

$$f(q, c, \ell, \varepsilon, |\bar{y}|) \cdot n^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8}. \quad \square$$

Now we obtain our main result Theorem 6 by simply applying Theorem 7 repeatedly, once for each quantifier. This will require no new insights, but will be a bit tedious to analyze. To help our inductive proof, we prove the following stronger statement. Then Theorem 6 follows as a special case when φ has no free variables, $p = q$ and $r = 0$.

Lemma 10. *Let \mathcal{C} be a monadically stable graph class admitting flip-closed sparse weak neighborhood covers with spread σ , and let $\rho = (6q\sigma(2^q) + 1)(2^q + 1)$. Given as input*

- *a (\mathcal{C}, ρ) -history $\mathcal{H} = (G_0, \mathcal{I}_0), \dots, (G_\ell, \mathcal{I}_\ell)$ of length ℓ ,*
- *a coloring G of G_ℓ with a signature of at most c colors,*
- *a formula $\varphi(\bar{y})$ with quantifier rank at most $p \leq q$,*
- *sets $W_1, \dots, W_{|\bar{y}|}$, each contained in an r -neighborhood of G , and*
- *an efficient $\text{MC}(\mathcal{C}, q, \ell + 1, c + 3)$ -algorithm,*

one can compute sets $U_1, \dots, U_t \subseteq V(G)$, for some constant t depending only on p, q, c , and $|\bar{y}|$, as well as a (U_1, \dots, U_t) -guarded formula $\xi(\bar{y})$ of quantifier rank p . Each U_i is contained in an $(r + q \cdot 3\sigma(2^q))$ -neighborhood of G and for all tuples $\bar{w} \in W_1 \times \dots \times W_{|\bar{y}|}$ we have

$$G \models \varphi(\bar{w}) \iff G\langle U_1, \dots, U_t \rangle \models \xi(\bar{w}).$$

There exists a function $f(p, q, c, \ell, \varepsilon)$ such that for every $\varepsilon > 0$, the run time of this procedure is bounded by

$$f(p, q, c, \ell, \varepsilon) \cdot |V(G)|^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8},$$

where $d := \text{game-depth}(\mathcal{C}, \rho) - \ell$ bounds the number of rounds needed to win the remaining Flipper game.

Proof. Let $\varepsilon > 0$. We prove the lemma by induction on p . For $p = 0$, note that every quantifier-free formula is \emptyset -guarded, and thus we can set $\xi(\bar{y}) := \varphi(\bar{y})$ and there is nothing more to show. Thus assume $p > 0$ and that the statement holds for $p - 1$. We will construct an algorithm for p using the assumed algorithm for $p - 1$ as a subroutine.

By normalization, $|\varphi|$ depends only on p, c and $|\bar{y}|$. Furthermore $\varphi(\bar{y})$ is a boolean combination of formulas of the form $\exists x \psi(\bar{y}, x)$ of quantifier rank at most p . Thus, it is sufficient to prove the theorem for a single such formula $\exists x \psi(\bar{y}, x)$. We apply Theorem 7 giving it as input

- the history $\mathcal{H} = (G_0, \mathcal{I}_0), \dots, (G_\ell, \mathcal{I}_\ell)$,
- the coloring G of G_ℓ with a signature of at most c colors,
- the formula $\exists x \psi(\bar{y}, x)$ of quantifier rank at most $p \leq q$,
- the sets $W_1, \dots, W_{|\bar{y}|}$, each contained in an r -neighborhood of G , and
- the given $\text{MC}(\mathcal{C}, q, \ell + 1, c + 3)$ -algorithm.

In time

$$f'(q, c, \ell, \varepsilon, |\bar{y}|) \cdot |V(G)|^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8} \tag{4}$$

this yields sets $R_1, \dots, R_{t'} \subseteq V(G)$ for some constant t' depending only on q, c , and $|\bar{y}|$. Each R_i is contained in an $(r + 3\sigma(2^q))$ -neighborhood of G , such that for all tuples $\bar{w} \in W_1 \times \dots \times W_{|\bar{y}|}$, we have

$$G \models \exists x \psi(\bar{w}, x) \iff \bigvee_{i=1}^{t'} G\langle R_i \rangle \models \exists x \in R_i \psi(\bar{w}, x). \tag{5}$$

For each $i \in [t']$ we apply the algorithm for $p - 1$ given by the induction hypothesis on

- the history \mathcal{H} , graph G , and $\text{MC}(\mathcal{C}, q, \ell + 1, c + 3)$ -algorithm,
- the formula $\psi(\bar{y}, x)$ of quantifier rank at most $p - 1 \leq q$,
- the sets $W_1, \dots, W_{|\bar{y}|}, R_i \subseteq V(G)$, each contained in an $(r + 3\sigma(2^q))$ -neighborhood of G .

In time

$$f(p - 1, q, c, \ell, \varepsilon, |\bar{y}| + 1) \cdot |V(G)|^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8} \quad (6)$$

this yields a family of guarding sets \mathcal{U}_i with $|\mathcal{U}_i|$ depending on $p - 1$, q , c , and $|\bar{y}|$, as well as a \mathcal{U}_i -guarded formula $\xi_i(\bar{y})$ of quantifier rank $q - 1$. Each $U \in \mathcal{U}_i$ is contained in an $(r + 3\sigma(2^q) + (q - 1) \cdot 3\sigma(2^{q-1}))$ -neighborhood of G (and thus in an $(r + q \cdot 3\sigma(2^q))$ -neighborhood of G). For all tuples $\bar{w}v \in W_1 \times \dots \times W_{|\bar{y}|} \times R_i$ we have

$$G \models \psi(\bar{w}, v) \iff G\langle \mathcal{U}_i \rangle \models \xi_i(\bar{w}, v).$$

Since the above statement holds no matter how $v \in R_i$ is chosen, existentially quantifying $v \in R_i$ preserves the equivalence. Hence, for all tuples $\bar{w} \in W_1 \times \dots \times W_{|\bar{y}|}$

$$G\langle R_i \rangle \models \exists x \in R_i \psi(\bar{w}, x) \iff G\langle R_i \rangle \langle \mathcal{U}_i \rangle \models \exists x \in R_i \xi_i(\bar{w}, x). \quad (7)$$

Combining (5) and (7) yields for every $\bar{w} \in W_1 \times \dots \times W_{|\bar{y}|}$,

$$G \models \exists x \psi(\bar{w}, x) \iff \bigvee_{i=1}^{t'} G\langle R_i \rangle \langle \mathcal{U}_i \rangle \models \exists x \in R_i \xi_i(\bar{w}, x),$$

which is equivalent to

$$G\langle R_1 \rangle \langle \mathcal{U}_1 \rangle \dots \langle R_{t'} \rangle \langle \mathcal{U}_{t'} \rangle \models \bigvee_{i=1}^{t'} \exists x \in R_i \xi_i(\bar{w}, x).$$

Thus, we can define our guarding sets $\mathcal{U} = \{U_1, \dots, U_t\}$ as $\mathcal{U} := \{R_1, \dots, R_{t'}\} \cup \bigcup_{i=1}^{t'} \mathcal{U}_i$.

The running time is bounded by the bound (4) for the invocation of Theorem 7, plus t' times the bound (6) for the recursive calls with $p - 1$, plus some minor bookkeeping overhead. We can choose $f(p, q, c, \ell, \varepsilon)$ such that this is at most

$$f(p, q, c, \ell, \varepsilon) \cdot |V(G)|^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8}. \quad \square$$

5.3 Reducing the Evaluation Radius

Our overall goal is to evaluate a sentence with quantifier rank q on a graph G . In the previous section, we have rewritten the sentence into an equivalent \mathcal{U} -guarded sentence of the same quantifier rank using guards $\mathcal{U} = \{U_1, \dots, U_t\}$. Each of the sets $U_i \subseteq V(G)$ is contained in an $r := q \cdot 3\sigma(2^q)$ -neighborhood of G and thus the induced graph $G[U_1 \cup \dots \cup U_t]$ consists of components, which are contained in neighborhoods with radius at most $(2r + 1)t$ in G .

One could imagine evaluating the \mathcal{U} -guarded sentence on G by recursing into each of these components and to compute flips using the strategy for the radius- $(2r + 1)t$ Flipper game. Let us argue that this cannot work. The radius of the Flipper game is not allowed to grow over time, since otherwise the game is not guaranteed to terminate in a fixed number of rounds. In our construction,

however, the number t of guards depends on the number of colors c added over time and thus grows with the number of rounds of the Flipper game played so far. Thus, we are not allowed to recurse into components with radius $(2r + 1)t$. We have to choose a fixed radius ρ for the Flipper game, depending only on q and \mathcal{C} .

In this section, we show that we can evaluate the \mathcal{U} -guarded sentence by only looking at subgraphs of G that are contained in neighborhoods of radius $\rho := (2r + 1)(2^q + 1)$, a quantity that depends only on q and \mathcal{C} and does not grow over time. This is a consequence of the following Theorem 8. Remember that, to avoid lengthy additional notation, \mathcal{U} will refer both to unary predicates guarding a formula, as well as the corresponding vertex sets in a graph G that interpret these predicates. It will be clear from the context which one is meant.

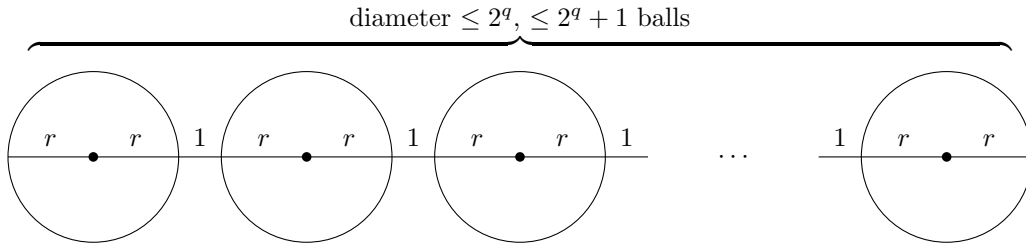
Theorem 8. *For a given \mathcal{U} -guarded sentence φ with quantifier rank at most q and symmetric relation $\mathcal{R} \subseteq \mathcal{U} \times \mathcal{U}$, one can compute a sentence $\varphi_{\mathcal{R}}$ such that for every graph G and set $\mathcal{U} \subseteq \mathcal{P}(V(G))$ satisfying $\mathcal{R} = \{(U, W) \in \mathcal{U} \times \mathcal{U} \mid U \text{ and } W \text{ share a vertex or a connecting edge in } G\}$, we have*

$$G(\mathcal{U}) \models \varphi \iff G(\mathcal{U}) \models \varphi_{\mathcal{R}}.$$

Moreover, $\varphi_{\mathcal{R}}$ is a boolean combination of sentences with quantifier rank at most q and each sentence mentioned in $\varphi_{\mathcal{R}}$ is \mathcal{U}' -guarded for some $\mathcal{U}' \subseteq \mathcal{U}$ such that the graph $(\mathcal{U}, \mathcal{R})[\mathcal{U}']$ has diameter at most 2^q .

In particular, if each set of $\mathcal{U} \subseteq \mathcal{P}(V(G))$ is contained in an r -neighborhood of G , then $\bigcup \mathcal{U}'$ is contained in a subgraph of G with diameter at most $(2r + 1)(2^q + 1)$.

To see that the final “In particular, ...” part follows from the central part of the statement, assume each set of $\mathcal{U} \subseteq \mathcal{P}(V(G))$ is contained in an r -neighborhood of G . For a \mathcal{U}' -guarded sentence ξ in $\varphi_{\mathcal{R}}$, the graph $(\mathcal{U}, \mathcal{R})[\mathcal{U}']$ has diameter at most 2^q and all $(U_1, U_2) \in \mathcal{R}$ share a vertex or a connecting edge in G . As can be seen in the figure below, $\bigcup \mathcal{U}'$ is contained in a subgraph of G with diameter at most $(2r + 1)(2^q + 1)$.



We prove the central part of Theorem 8 inductively using a stronger statement involving formulas with free variables.

Lemma 11. *For a given \mathcal{U} -guarded formula $\varphi(\bar{y})$ with quantifier rank at most q , symmetric relation $\mathcal{R} \subseteq \mathcal{U} \times \mathcal{U}$ and sequence $U_1, \dots, U_{|\bar{y}|} \in \mathcal{U}$ one can compute a formula $\varphi_{\mathcal{R}}(\bar{y})$ such that for every graph G , set $\mathcal{U} \subseteq \mathcal{P}(V(G))$ associated with the predicates \mathcal{U} satisfying $\mathcal{R} = \{(U, W) \in \mathcal{U} \times \mathcal{U} \mid U \text{ and } W \text{ share a vertex or a connecting edge in } G\}$, and every $\bar{w} \in U_1 \times \dots \times U_{|\bar{y}|}$ we have*

$$G(\mathcal{U}) \models \varphi(\bar{w}) \iff G(\mathcal{U}) \models \varphi_{\mathcal{R}}(\bar{w}).$$

Moreover, $\varphi_{\mathcal{R}}(\bar{y})$ is a boolean combination of formulas with quantifier rank at most q and for each formula ξ mentioned in $\varphi_{\mathcal{R}}$ there exists $\mathcal{U}' \subseteq \mathcal{U}$ such that ξ is \mathcal{U}' -guarded, $(\mathcal{U}, \mathcal{R})[\mathcal{U}']$ has diameter at most 2^q , and $\{U_i \mid y_i \in \text{free}(\xi)\} \subseteq \mathcal{U}'$.

Proof. We consider an arbitrary graph G and $\mathcal{U} \subseteq \mathcal{P}(V(G))$ associated with the predicates \mathcal{U} such that for all $U, W \in \mathcal{U}$ we have $(U, W) \in \mathcal{R}$ if and only if U, W share a vertex or a connecting edge in G . Let us also fix a sequence $U_1, \dots, U_{|\bar{y}|} \in \mathcal{U}$. We prove the claim by induction over the structure of φ .

Atoms. Since φ is an atom, it is \emptyset -guarded and has either one or two free variables. Assume $\varphi(y_1)$ has a single free variable. We set $\varphi_{\mathcal{R}} := \varphi$ and $\mathcal{U}' := \{U_1\}$. Then $\varphi_{\mathcal{R}}$ itself is \mathcal{U}' -guarded and $(\mathcal{U}, \mathcal{R})[\mathcal{U}']$ trivially has diameter $0 \leq 2^0$. Assume now $\varphi(y_1, y_2)$ is a binary atom, that is, without loss of generality either $E(y_1, y_2)$ or $(y_1 = y_2)$. If $(U_1, U_2) \in \mathcal{R}$ we set $\varphi_{\mathcal{R}} := \varphi$ and $\mathcal{U}' := \{U_1, U_2\}$. Again, $\varphi_{\mathcal{R}}$ is \mathcal{U}' -guarded and $(\mathcal{U}, \mathcal{R})[\mathcal{U}']$ has diameter $1 = 2^0$. Otherwise, $(U_1, U_2) \notin \mathcal{R}$ and U_1, U_2 neither share a vertex nor a connecting edge in G . This implies $G \not\models E(w_1, w_2)$ and $G \not\models (w_1 = w_2)$ for all $w_1 \in U_x, w_2 \in U_y$. We set $\varphi_{\mathcal{R}}$ to be the false atom \perp and $\mathcal{U}' = \emptyset$.

Boolean Combinations. If φ is of the form $\psi^1 \wedge \psi^2$ or $\neg \psi^1$ the construction is obvious: We obtain $\varphi_{\mathcal{R}}^1$ and $\varphi_{\mathcal{R}}^2$ via induction and set either $\varphi_{\mathcal{R}} := \psi_{\mathcal{R}}^1 \wedge \psi_{\mathcal{R}}^2$ or $\varphi_{\mathcal{R}} := \neg \psi_{\mathcal{R}}^1$.

Existential Quantifiers. Assume $\varphi(\bar{y}) = \exists x \in U \psi(\bar{y}x)$. We apply the statement inductively on $\psi(\bar{y}x)$ (extending the sequence $U_1, \dots, U_{|\bar{y}|}$ with U) and obtain a boolean combination $\psi_{\mathcal{R}}(\bar{y}x)$ of formulas with quantifier rank at most $q - 1$ such that for every $\bar{w}v \in U_{y_1} \times \dots \times U_{y_{|\bar{y}|}} \times U$

$$G\langle \mathcal{U} \rangle \models \psi(\bar{w}v) \iff G\langle \mathcal{U} \rangle \models \psi_{\mathcal{R}}(\bar{w}v).$$

For each formula ξ mentioned in $\psi_{\mathcal{R}}(\bar{y}x)$ we have, by induction, a set $\mathcal{U}_{\xi} \subseteq \mathcal{U}$ such that ξ is \mathcal{U}_{ξ} -guarded and $(\mathcal{U}, \mathcal{R})[\mathcal{U}_{\xi}]$ has diameter at most 2^{q-1} . The additional crucial property we obtain by induction is that $x \in \text{free}(\xi)$ implies $U \in \mathcal{U}_{\xi}$. We partition the formulas mentioned in $\psi_{\mathcal{R}}(\bar{y}x)$ into sets Ψ_x and $\Psi_{\#}$, where Ψ_x contains all formulas ξ with $x \in \text{free}(\xi)$, and $\Psi_{\#}$ contains all ξ with $x \notin \text{free}(\xi)$. The formulas in $\Psi_{\#}$ are independent of x and we can thus write

$$\exists x \in U \psi_{\mathcal{R}}(\bar{y}x) \equiv \bigvee_{t: \Psi_{\#} \rightarrow \{\perp, \top\}} \left(\left(\bigwedge_{\xi \in \Psi_{\#}} (\xi(\bar{y}) \leftrightarrow t(\xi)) \right) \wedge \exists x \in U \psi_{\mathcal{R}}(\bar{y}x) \right).$$

Now on the right-hand side, every occurrence of $\psi_{\mathcal{R}}$ is in a scope where the truth value of every $\xi \in \Psi_{\#}$ is determined. Thus, we can replace every occurrence of ξ in $\psi_{\mathcal{R}}$ with said truth value $t(\xi) \in \{\perp, \top\}$. Let $\psi_{\mathcal{R}}^t$ be the formula obtained from $\psi_{\mathcal{R}}$ by replacing each occurrence of $\xi \in \Psi_{\#}$ with $t(\xi)$. We obtain the equivalence

$$\exists x \in U \psi_{\mathcal{R}}(\bar{y}x) \equiv \varphi_{\mathcal{R}}(\bar{y}) := \bigvee_{t: \Psi_{\#} \rightarrow \{\perp, \top\}} \left(\left(\bigwedge_{\xi \in \Psi_{\#}} (\xi(\bar{y}) \leftrightarrow t(\xi)) \right) \wedge \exists x \in U \psi_{\mathcal{R}}^t(\bar{y}x) \right).$$

Hence, for every $\bar{w} \in U_{y_1} \times \dots \times U_{y_{|\bar{y}|}}$

$$G\langle \mathcal{U} \rangle \models \varphi(\bar{w}) \iff G\langle \mathcal{U} \rangle \models \varphi_{\mathcal{R}}(\bar{w}).$$

We observe that $\varphi_{\mathcal{R}}(\bar{y})$ is a boolean combination of old formulas from $\Psi_{\#}$ and new formulas of the form $\exists x \in U \psi_{\mathcal{R}}^t$. All these formulas have quantifier rank at most q .

Consider now a new formula $\omega := \exists x \in U \psi_{\mathcal{R}}^t$ and let $\mathcal{U}_{\omega} := \bigcup_{\xi \in \Psi_x} \mathcal{U}_{\xi}$. Since $\psi_{\mathcal{R}}^t$ eliminated all formulas from $\Psi_{\#}$, we know that ω is \mathcal{U}_{ω} -guarded. For all $\xi \in \Psi_x$ we have $x \in \text{free}(\xi)$ and thus, as noted previously, $U \in \mathcal{U}_{\xi}$. By induction, each graph $(\mathcal{U}, \mathcal{R})[\mathcal{U}_{\xi}]$ has diameter at most 2^{q-1} . This means $(\mathcal{U}, \mathcal{R})[\mathcal{U}_{\omega}]$ is covered by graphs that all overlap in U and have diameter at most 2^{q-1} , implying that $(\mathcal{U}, \mathcal{R})[\mathcal{U}_{\omega}]$ has diameter at most 2^q . Since $\text{free}(\omega) \subseteq \bigcup_{\xi \in \Psi_x} \text{free}(\xi)$, we also have $\{U_i \mid y_i \in \text{free}(\omega)\} \subseteq \mathcal{U}_{\omega}$. \square

We remark that one obtains the classical Feferman–Vaught theorem [Fef57, Mak04] for disjoint unions as a corollary of Theorem 8: Assume one wants to evaluate a sentence on the disjoint union of graphs G_1 and G_2 . Replace every quantifier $\exists x \psi$ with $\exists x \in V(G_1) \psi \vee \exists x \in V(G_2) \psi$, and proceed similarly for all universal quantifiers. This gives us an equivalent \mathcal{U} -guarded sentence with $\mathcal{U} = \{V(G_1), V(G_2)\}$. For the relation $\mathcal{R} \subseteq \mathcal{U} \times \mathcal{U}$ corresponding to the disjoint union of G_1 and G_2 , the graph $(\mathcal{U}, \mathcal{R})$ consists of two vertices $V(G_1)$ and $V(G_2)$ that are not connected by an edge. We apply Theorem 8 with relation \mathcal{R} . Each sentence in the boolean combination we obtain is \mathcal{U}' -guarded for some $\mathcal{U}' \subseteq \mathcal{U}$ such that $(\mathcal{U}, \mathcal{R})[\mathcal{U}']$ is connected. This leaves only $\mathcal{U}' \subseteq \{V(G_1)\}$ and $\mathcal{U}' \subseteq \{V(G_2)\}$. Hence, we have a boolean combination of sentences that are evaluated in either G_1 or G_2 .

5.4 Main Result

We are ready to prove the main result. We start by combining the observations from the previous two subsections into an inductive step on the depth of the Flipper game.

Theorem 9. *Let \mathcal{C} be a monadically stable graph class admitting flip-closed sparse weak neighborhood covers with spread σ , and let $q, \ell \in \mathbb{N}$. If \mathcal{C} for every $c \in \mathbb{N}$ has an efficient $\text{MC}(\mathcal{C}, q, \ell + 1, c)$ -algorithm, then \mathcal{C} for every $c \in \mathbb{N}$ also has an efficient $\text{MC}(\mathcal{C}, q, \ell, c)$ -algorithm.*

Proof. Let $\rho := (6q\sigma(2^q) + 1)(2^q + 1)$. The $\text{MC}(\mathcal{C}, q, \ell, c)$ -algorithm we will construct gets as input a (\mathcal{C}, ρ) -history $(G_0, \mathcal{I}_0), \dots, (G_\ell, \mathcal{I}_\ell)$, a coloring G of G_ℓ with at most c colors, and a sentence φ with quantifier rank q . At first, we call Theorem 6. This gives us a \mathcal{U} -guarded sentence ξ of quantifier rank q such that

$$G \models \varphi \iff G\langle \mathcal{U} \rangle \models \xi. \quad (8)$$

Here, $\mathcal{U} \subseteq \mathcal{P}(V(G))$ is a set with $|\mathcal{U}|$ depending only on q, c , such that each $U \in \mathcal{U}$ is contained in an $(q \cdot 3\sigma(2^q))$ -neighborhood of G .

In time $\mathcal{O}(|\mathcal{U}|^2 \cdot |V(G)|^2)$, compute the relation $\mathcal{R} := \{(U, W) \in \mathcal{U} \times \mathcal{U} \mid U \text{ and } W \text{ share a vertex or a connecting edge in } G\}$. Next, we invoke Theorem 8. This gives us a boolean combination ξ^* of sentences ξ_1, \dots, ξ_k with quantifier rank at most q such that

$$G\langle \mathcal{U} \rangle \models \xi \iff G\langle \mathcal{U} \rangle \models \xi^*. \quad (9)$$

Each sentence ξ_i is \mathcal{U}_i -guarded for some $\mathcal{U}_i \subseteq \mathcal{U}$ such that $\bigcup \mathcal{U}_i$ is contained in a subgraph of G with diameter at most $\rho = (6q\sigma(2^q) + 1)(2^q + 1)$, and thus also in a ρ -neighborhood of G . The running time of Theorem 8 is insignificant compared to the running time of Theorem 6. The time spent so far is bounded by

$$f(q, c, \ell, \varepsilon) \cdot |V(G)|^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8} \quad (10)$$

for some function $f(q, c, \ell, \varepsilon)$ and every $\varepsilon > 0$, where $d := \text{game-depth}(\mathcal{C}, \rho) - \ell$ bounds the number of rounds needed to win the remaining Flipper game.

Now for every $i \in [k]$, we proceed similarly as in the paragraph *Flip and Type Computation* of Theorem 7 to decide whether $G\langle \mathcal{U} \rangle \models \xi_i$. Let $H_{\ell+1} := G[\bigcup \mathcal{U}_i]$ and as ξ_i is \mathcal{U}_i -guarded, we have

$$G\langle \mathcal{U} \rangle \models \xi_i \iff H_{\ell+1}\langle \mathcal{U} \rangle \models \xi_i. \quad (11)$$

Note that since $\bigcup \mathcal{U}_i$ is contained in a ρ -neighborhood of G , the restriction to $H_{\ell+1}$ corresponds to a Connector move in the radius- ρ Flipper game. We apply the radius- ρ Flipper strategy flip^* (for the class $\mathcal{C} \ni G_0$) to the graph $H_{\ell+1}$ and internal state \mathcal{I}_ℓ , yielding a flip F specified by flip sets A, B and a new internal state $\mathcal{I}_{\ell+1}$. By Theorem 2, this takes time

$$g_2(q) \cdot |V(G_0)|^2, \quad (12)$$

for some function $g_2(q)$.

Let $G_{\ell+1} := H_{\ell+1} \oplus \mathbf{F}$ and $G_{\ell+1}^+ = G_{\ell+1} \langle \mathcal{U} \rangle \langle A, B \rangle$. We construct ξ'_i from ξ_i by substituting every occurrence of the edge relation $E(x, y)$ with $E(x, y) \oplus ((x \in A \wedge y \in B) \vee (x \in B \wedge y \in A))$. Then

$$H_{\ell+1} \langle \mathcal{U} \rangle \models \xi_i \iff G_{\ell+1}^+ \models \xi'_i. \quad (13)$$

We can now extend \mathcal{H} to a (\mathcal{C}, ρ) -history of length $\ell + 1$ by appending the new pair $(G_{\ell+1}, \mathcal{I}_{\ell+1})$. We use the given efficient $\text{MC}(\mathcal{C}, q, \ell + 1, c + 2 + |\mathcal{U}|)$ -algorithm to decide in time

$$f_{\text{MC}}(q, \ell + 1, c + 2 + |\mathcal{U}|, \varepsilon) \cdot |V(G_{\ell+1}^+)|^{((1+\varepsilon)^{d-1})} \cdot |V(G_0)|^{9.8} \quad (14)$$

whether $G_{\ell+1}^+ \models \xi'_i$. By (11) and (13), this decides whether $G \langle \mathcal{U} \rangle \models \xi_i$.

Since we decided $G \langle \mathcal{U} \rangle \models \xi_i$ for all i , we can plug the truth values into the boolean combination ξ^* , telling us the answer to whether $G \langle \mathcal{U} \rangle \models \xi^*$. By (8) and (9), this finally gives us the answer whether $G \models \varphi$.

The total running time is bounded by (10) plus k times (12) and (14). Since both k and $|\mathcal{U}|$ are bounded by a function of q and c , we can choose $f_{\text{MC}}(q, \ell, c, \varepsilon)$ such that for every $\varepsilon > 0$, the total running time is bounded by

$$f_{\text{MC}}(q, \ell, c, \varepsilon) \cdot |V(G)|^{((1+\varepsilon)^d)} \cdot |V(G_0)|^{9.8}. \quad \square$$

We are ready to prove the stepping stone theorem. We control the run time by choosing $\varepsilon > 0$ as a function of the game depth.

Theorem 5. *Let \mathcal{C} be a monadically stable graph class admitting flip-closed sparse weak neighborhood covers. There exists a function $f: \mathbb{N} \rightarrow \mathbb{N}$ such that, given a graph $G \in \mathcal{C}$ and sentence φ , one can decide whether $G \models \varphi$ in time $f(|\varphi|) \cdot |V(G)|^{11}$.*

Proof. Let q be the quantifier rank of φ and $\rho := (6q\sigma(2^q) + 1)(2^q + 1)$. Recall that $\text{game-depth}(\mathcal{C}, \rho)$ bounds the number of rounds the Flipper strategy flip^* needs to win the Flipper game, that is, until a graph is reached that consists only of a single vertex. On such graphs, model checking is trivial. Hence, \mathcal{C} for every $c \in \mathbb{N}$ has an efficient $\text{MC}(\mathcal{C}, q, \text{game-depth}(\mathcal{C}, \rho), c)$ -algorithm. By repeated application of Theorem 9, \mathcal{C} also has an efficient $\text{MC}(\mathcal{C}, q, 0, |\Sigma|)$ -algorithm, where Σ is the signature of \mathcal{C} . Such an $\text{MC}(\mathcal{C}, q, 0, |\Sigma|)$ -algorithm can decide for every graph $G \in \mathcal{C}$ whether φ holds, and thus solves our problem. The running time is bounded by

$$f_{\text{MC}}(q, 0, |\Sigma|, \varepsilon) \cdot |V(G)|^{((1+\varepsilon)^{\text{game-depth}(\mathcal{C}, \rho)})} \cdot |V(G)|^{9.8}.$$

By choosing $\varepsilon := 1.2^{1/\text{game-depth}(\mathcal{C}, \rho)} - 1 > 0$, we get a running time of

$$f_{\text{MC}}(q, 0, |\Sigma|, \varepsilon) \cdot |V(G)|^{1.2} \cdot |V(G)|^{9.8} \leq f_{\text{MC}}(q, 0, |\Sigma|, \varepsilon) \cdot |V(G)|^{11}.$$

As our choice of ε and all other parameters of f_{MC} depends only on $|\varphi|$ and \mathcal{C} , we can choose for every class \mathcal{C} a function $f(|\varphi|)$ such that the runtime is bounded by $f(|\varphi|) \cdot |V(G)|^{11}$. \square

5.5 A Note on the Computability of $f(|\varphi|)$

In the previous theorems, we have proven a run time bound depending on a function $f(|\varphi|)$. It should be clear that $f(|\varphi|)$ is an incredibly fast-growing function, and it has to be, as shown in [FG04]. But let us stress that $f(|\varphi|)$ may in fact not even be computable. This is because in monadically stable or structurally nowhere dense classes the functions bounding the order of

transducible half graphs or the size of shallow clique minors may not be computable. This implies that the depth of the Flipper game $\text{game-depth}(\mathcal{C}, \rho)$ may not be computable, which in turn is a lower bound for $f(|\varphi|)$. Following [GKS17], one may define *effectively nowhere dense* classes where one requires the size of r -shallow clique minors to be bounded by a *computable* function of r . Similarly, in *effectively monadically stable classes*, the order of largest transducible half graph is bounded by a computable function of the transduction. Under the promise that the graph class of interest is *effective*, we believe that revisiting all underlying proofs from [DGK⁺22a, DMST22, GMM⁺23] in fact gives us a computable function $f(|\varphi|)$.

6 Approximating Weak Neighborhood Covers

Sparse neighborhood covers have been a central tool in the design of model checking algorithms on nowhere dense graphs, where they have been computed using mechanisms that are deeply tied to the sparse structure of the input graph. In this section, we observe that these ties can be cut, and in fact sparse weak neighborhood covers can be computed in *any* graph where they exist. While finding for a fixed spread s an r -neighborhood cover whose degree is minimal is NP-complete (with $s = 1$, $r = 0$ via a simple reduction from minimum membership set cover [KvRW⁺05]), we nevertheless find a $\mathcal{O}(\log(n)^2 + 1)$ -approximation in polynomial time. In our context, this is good enough, as logarithmic factors are insignificant (see Definition 4). Note that the running time of Theorem 10 is independent of r and s .

Theorem 10. *There is an algorithm that gets as input an n -vertex graph G and numbers $r, s \in \mathbb{N}$ and computes in time $\mathcal{O}(n^{9.8})$ a weak r -neighborhood cover with degree $\mathcal{O}(\log(n)^2 + 1)d^*$ and spread s , where d^* is the smallest number such that G admits a weak r -neighborhood cover with degree d^* and spread s .*

6.1 Linear Programming

In this section, we use randomized rounding to approximate neighborhood covers. For a given matrix $A \in \mathbb{R}^{M \times N}$ and vectors $b \in \mathbb{R}^M$, $c \in \mathbb{R}^N$, the corresponding *integer linear program (ILP)* asks for a vector $x \in \mathbb{R}^N$ maximizing or minimizing $c^T x$ under the constraint that $Ax \geq b$ and $x \in \mathbb{Z}^N$. While solving ILPs is NP-complete, one can solve the corresponding *linear programming (LP) relaxation*, obtained by removing the integrality constraint $x \in \mathbb{Z}^N$ in polynomial time. The idea behind randomized rounding [RT87] is to express the problem of choice as an ILP, solve the corresponding LP relaxation and to then round the fractional solution to get an approximate solution to the original problem. This method has recently been used by Dvořák to compactly represent short distances [Dvo22]. For more background to this technique we refer, for example, to Chapter 14 *Rounding Applied to Set Cover* of Vazirani’s book on approximation algorithms [Vaz10].

The running time of this approach is primarily dominated by the time it takes to solve an LP. While the best running time bounds approach matrix multiplication time [CLS21], for simplicity, we use the classical algorithm by Vaidya [Vai89] with a running time of $\mathcal{O}((N+M)^{1.5}NL)$, where N is the number of variables, M is the number of constraints, and L roughly equals the number of bits in the input. The precise definition of L is a bit cumbersome: For a given LP with matrix $A \in \mathbb{Z}^{M \times N}$ and vectors $b \in \mathbb{Z}^M$, $c \in \mathbb{Z}^N$, the factor L is defined as

$$L := \log(1 + \det_{\max}) + \log(N + M),$$

where \det_{\max} denotes the largest absolute value of the determinant of any square submatrix of

$$\begin{pmatrix} c^T & 0 \\ A & b \end{pmatrix}.$$

If all entries in A, b, c are integers from $\{-1, 0, 1\}$, the running time bound can be simplified. With S_n being the set of all permutations on $\{1, \dots, n\}$, we can bound for every $n \times n$ submatrix A' of the above matrix

$$\det(A') := \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^n A'_{i, \sigma(i)} \leq |S_n| \leq n! \leq n^n.$$

This gives us with $n \leq N + M$,

$$L \leq \log(1 + (N + M)^{N+M}) + \log(N + M) \leq \mathcal{O}((N + M) \cdot \log(N + M)) \leq \mathcal{O}((N + M)^{1.1}).$$

Thus, Vaidya's algorithm has a running time of $\mathcal{O}((N + M)^{2.6} \cdot N)$.

6.2 ILP Formulation

We believe the central contribution of this section is the following robust ILP formulation for weak r -neighborhood covers. Let G be a graph and let \mathcal{X} be a weak r -neighborhood cover of G with spread s . Recall that for every $X \in \mathcal{X}$ there exists a vertex $\text{center}(X) \in V(G)$ with $X \subseteq N_s[\text{center}(X)]$. By merging all clusters with the same center, we can assume that there exists a set $S \subseteq V(G)$, such that $\mathcal{X} = \{X_w \mid w \in S\}$. We call X_w the *cluster around w* .

In order to compute a weak r -neighborhood cover with minimal degree and spread s in a graph G , we denote by $\text{cover-ILP}(G, r, s)$ the following ILP with variable $d \in \mathbb{Z}$ and binary variables $p_{vw}, q_{uw} \in \{0, 1\}$ for $u, v, w \in V(G)$ with the following intuitive meaning. We want $p_{vw} \geq 1$ whenever $N_r[v]$ is contained in the cluster around w , that is, $N_r[v] \subseteq X_w$, and $q_{uw} \geq 1$ whenever u is contained in the cluster around w , that is, $u \in X_w$. This intuition leads to the following ILP.

minimize d such that

$$\forall v \in V(G) \quad \sum_{w: N_r[v] \subseteq N_s[w]} p_{vw} \geq 1 \tag{1}$$

$$\forall u \in V(G) \quad \sum_w q_{uw} \leq d \tag{2}$$

$$\forall v, w \in V(G) \quad \forall u \in N_r[v] \quad q_{uw} \geq p_{vw} \tag{3}$$

$$\forall w \in V(G) \quad \forall u \notin N_s[w] \quad q_{uw} = 0 \tag{4}$$

Equations (1) enforce that for all v , the neighborhood $N_r[v]$ is contained in a cluster centered around a nearby vertex w . Similarly, (2) enforces that each v is contained in at most d clusters centered around some vertex w . The crucial equation (3) relates the variables q_{uw} and p_{vw} by stating that if a cluster around w contains $N_r[v]$, then it also contains u for all $u \in N_r[v]$.

Equations (4) ensure that the spread of the resulting cover is bounded by s . They are however not needed and only listed for the sake of presentation: no p_{vw} with $u \in N_r[v]$ and $u \notin N_s[w]$ is subject to an equation from (1), as we have $N_r[v] \not\subseteq N_s[w]$. We can therefore set $p_{vw} = 0$, and hence also $q_{uw} = 0$, without violating a constraint or increasing the cost.

Lemma 12. *If a graph G has a weak r -neighborhood cover with degree d and spread s , then $\text{cover-ILP}(G, r, s)$ has a solution of value at most d .*

Proof. Assume G has a weak r -neighborhood cover \mathcal{X} with degree d and spread s . As discussed above, we may assume that there exists a set $S \subseteq V(G)$ such that $\mathcal{X} = \{X_w \mid w \in S\}$ with $X_w \subseteq N_s[w]$. Set $p_{vw} = 1 \iff N_r[v] \subseteq X_w$ and $q_{uw} = 1 \iff u \in X_w$. If $N_r[v] \subseteq X_w$ then $N_r[v] \subseteq N_s[w]$ and thus (1) is satisfied. The degree bound d of \mathcal{X} implies that (2) is satisfied. Since \mathcal{X} has spread s , we have $u \notin X_w$ if $u \notin N_s[w]$ and consequently $q_{uw} = 0$, satisfying (4). At last, (3) simply states that for all $v, w \in V(G)$, and $u \in N_r[v]$ with $N_r[v] \subseteq X_w$ we have $u \in X_w$. Thus, (3) is also true. We conclude that the ILP has a solution with value at most d . \square

The *LP relaxation* of an ILP is obtained by removing all integrality constraints. For the ILP above this means one allows $d \in \mathbb{R}$, and $0 \leq p_{vw}, q_{uw} \leq 1$ instead of just $p_{vw}, q_{uw} \in \{0, 1\}$ for all $u, v, w \in V(G)$. The matrix A and vectors b, c representing this LP only contain entries from $\{-1, 0, 1\}$. As discussed above, such an LP can be solved using Vaidya's algorithm [Vai89] in time $\mathcal{O}((N + M)^{2.6}N)$. Here, $N \in \mathcal{O}(|V(G)|^2)$ and $M \in \mathcal{O}(|V(G)|^3)$, proving the following lemma.

Lemma 13. *One can solve the LP relaxation of $\text{cover-ILP}(G, r, s)$ in time $\mathcal{O}(|V(G)|^{9.8})$.*

6.3 Fractional Weak Neighborhood Covers

Since the LP relaxation is less restrictive than the original ILP and by Lemma 12, the solution given by Lemma 13 is at least as good as the minimum d such that G has a weak r -neighborhood cover with degree d and spread s . Next, we convert the solution to the LP relaxation into a so-called *fractional weak r -neighborhood cover* defined as follows.

Definition 5. A *fractional weak r -neighborhood cover* with degree d and spread s of a graph G is a family \mathcal{X} of subsets of $V(G)$, called *clusters*, together with a function $f : \mathcal{X} \rightarrow [0, 1]$ (which can be thought of as probability assignments) such that

- every cluster is a subset of an s -neighborhood in G ,
- for every vertex v , the probabilities of the clusters in which the r -neighborhood around v is fully contained sum up to at least 1, that is,

$$\sum_{X \in \mathcal{X} : N_r[v] \subseteq X} f(X) \geq 1,$$

- for every vertex, the probabilities of the clusters containing it sum up to at most d , that is,

$$\sum_{X \in \mathcal{X} : v \in X} f(X) \leq d.$$

Note that every fractional weak r -neighborhood cover with integer probabilities $f : \mathcal{X} \rightarrow \{0, 1\}$ is also a weak r -neighborhood cover.

Lemma 14. *Every solution to the LP relaxation of $\text{cover-ILP}(G, r, s)$ with degree d can be converted in time $\mathcal{O}(|V(G)|^2)$ into a fractional weak r -neighborhood cover with degree $d + 1$ and spread s .*

Proof. Let the solution to the LP relaxation be $0 \leq p_{vw}, q_{uw} \leq 1$ for $u, v, w \in V(G)$ and let $|V(G)| = n$. We define $P_{vw}, Q_{uw} \in [n]$ as $P_{vw} := \lceil n \cdot p_{vw} \rceil$ and $Q_{uw} := \lceil n \cdot q_{uw} \rceil$. Then

$$\forall v \in V(G) \quad \sum_{w: N_r[v] \subseteq N_s[w]} P_{vw} \geq n \quad (1)$$

$$\forall u \in V(G) \quad \sum_w Q_{uw} \leq \sum_w (1 + n \cdot q_{vw}) \leq (d+1) \cdot n \quad (2)$$

$$\forall v, w \in V(G) \quad \forall u \in N_r[v] \quad Q_{uw} \geq P_{vw} \quad (3)$$

$$\forall w \in V(G) \quad \forall u \notin N_s[w] \quad Q_{uw} = 0 \quad (4)$$

We build the desired fractional cover \mathcal{X} by creating for every $w \in V(G)$ and $i \in [n]$ a cluster $X_w^i := \{u \in V(G) \mid Q_{uw} \geq i\}$ with probability $f(X_w^i) = 1/n$. Using (4), the spread of \mathcal{X} is clearly bounded by s . For every vertex u observe that $\{X \in \mathcal{X} : u \in X\} = \{X_w^i : w \in V(G) \wedge Q_{uw} \geq i\}$ and therefore

$$|\{X \in \mathcal{X} : u \in X\}| = \sum_w Q_{uw} \quad (5)$$

Thus, the degree of u in \mathcal{X} is bounded by

$$\sum_{X \in \mathcal{X} : u \in X} f(X) = \sum_{X \in \mathcal{X} : u \in X} 1/n \stackrel{\text{by (5)}}{=} \sum_w Q_{uw}/n \stackrel{\text{by (2)}}{\leq} (d+1).$$

Let $i \leq P_{vw}$. Then by (3) for all $u \in N_r[v]$ we have $i \leq Q_{uw}$, and thus $u \in X_w^i$. Thus, for all $1 \leq i \leq P_{vw}$ we have $N_r[v] \subseteq X_w^i$. Therefore

$$\sum_{X \in \mathcal{X} : N_r[v] \subseteq X} f(X) = \sum_{X \in \mathcal{X} : N_r[v] \subseteq X} 1/n \geq \sum_{w \in N_s[v]} P_{vw}/n \stackrel{\text{by (1)}}{\geq} 1. \quad \square$$

6.4 Randomized Sampling

The previous lemmas together let us compute a good fractional weak r -neighborhood cover (\mathcal{X}, f) of G . By simply sampling clusters X from \mathcal{X} with probability $24 \ln(|V(G)|) f(X)$, the following Lemma 15 proves the existence of a (non-fractional) weak r -neighborhood cover whose degree is at most a factor $36 \ln(|V(G)|)$ larger than the optimum. The argument can be easily turned into a randomized algorithm with high success probability. However, we ultimately want a fully deterministic algorithm. One may derandomize the algorithm using the method of conditional probabilities and an adequate pessimistic estimator, but the details of this process are a bit tedious. Instead, we use the purely existential Lemma 15 to reduce the search space for our weak r -neighborhood cover from the exponential number of all subsets of s -neighborhoods in G to the at most $\mathcal{O}(|V(G)|^2)$ clusters from \mathcal{X} . After the reduction of the search space, we will compute the final weak r -neighborhood cover by a simple reduction to a set cover variant called *Minimum Membership Set Cover*, which can be efficiently approximated.

Lemma 15. *Assume a graph G with $n \geq 5$ vertices has a fractional weak r -neighborhood cover (\mathcal{X}, f) with degree d and spread s . Then G has a weak r -neighborhood cover $\mathcal{Y} \subseteq \mathcal{X}$ with degree $36 \ln(n)d$ and spread s .*

Proof. We use the probabilistic method. Let us sample a multiset \mathcal{Y} of clusters from \mathcal{X} by sampling each $X \in \mathcal{X}$ independently $24 \ln(n)$ many times with probability $f(X)$. Note that \mathcal{Y} is a *multiset* in the sense that it can contain multiple copies (up to $24 \ln(n)$ many) of each cover from \mathcal{X} . For every vertex v , let A_v be the number of clusters in \mathcal{Y} fully containing the r -neighborhood around v . By linearity of expectation,

$$E[A_v] = 24 \ln(n) \cdot \sum_{X \in \mathcal{X}: N_r[v] \subseteq X} f(X) \geq 24 \ln(n).$$

Similarly, let B_v be the number of clusters in \mathcal{Y} that contain v . We argue as above that

$$E[A_v] \leq E[B_v] = 24 \ln(n) \cdot \sum_{X \in \mathcal{X}: v \in X} f(X) \leq 24 \ln(n)d.$$

Since A_v is a sum of independent binary variables, the Chernoff bound⁴ yields

$$P(|A_v - E[A_v]| \geq E[A_v]/2) \leq 2e^{-E[A_v]/12} \leq 2e^{-24 \ln(n)/12} = 2(e^{\ln(n)})^{-2} = 2/n^2.$$

Using $E[B_v] \geq E[A_v]$ we have

$$P(|B_v - E[B_v]| \geq E[B_v]/2) \leq 2e^{-E[B_v]/12} \leq 2e^{-E[A_v]/12} \leq 2/n^2.$$

By the union bound, the probability that for some v either $|A_v - E[A_v]| \geq E[A_v]/2$ or $|B_v - E[B_v]| \geq E[B_v]/2$ is (with $n \geq 5$) at most $\sum_{v \in V(G)} (2/n^2 + 2/n^2) = 4/n < 1$. Hence, there exists \mathcal{Y} such that $|A_v - E[A_v]| \leq E[A_v]/2$ and $|B_v - E[B_v]| \leq E[B_v]/2$ for all $v \in V(G)$. This implies $A_v \geq \frac{1}{2}E[A_v] \geq 12 \ln(n) \geq 1$ and $B_v \leq \frac{3}{2}E[B_v] \leq 36 \ln(n)d$ for all $v \in V(G)$, proving the statement. \square

6.5 Minimum Membership Set Cover Reduction

We have computed a good fractional weak r -neighborhood cover (\mathcal{X}, f) and know that a sufficiently good (non-fractional) weak r -neighborhood cover can be chosen as a subset of \mathcal{X} . Hence, to find our solution it is sufficient to choose a subset $\mathcal{Y} \subseteq \mathcal{X}$ covering all r -neighborhoods of G such that the maximal number of sets a vertex is contained in is as small as possible.

As we will see soon, the remaining difficulty is captured by the following *Minimum Membership Set Cover* (MMSC) problem. Let V be our universe and \mathcal{X} be a collection of subsets of V such that $\bigcup_{X \in \mathcal{X}} X = V$. The MMSC problem asks for a subset $\mathcal{Y} \subseteq \mathcal{X}$ covering all elements in V such that the maximal number of times a vertex is covered $M(V, \mathcal{Y}) := \max_{v \in V} |\{X \in \mathcal{Y} \mid v \in X\}|$ is minimal. While the problem is NP-hard, it can be approximated within a logarithmic factor.

Lemma 16 (Theorem 3 of [KvRW⁺05]). *For any MMSC instance (V, \mathcal{X}) , one can compute in time $\mathcal{O}((|V| + |\mathcal{X}|)^{2.6} |\mathcal{X}|)$ a set $\mathcal{Y} \subseteq \mathcal{X}$ with*

$$M(V, \mathcal{Y}) \leq (1 + \mathcal{O}(1/\sqrt{z}))(\log(|V|) + 1)z,$$

where

$$z = \min\{M(V, \mathcal{Y}^*) \mid \mathcal{Y}^* \subseteq \mathcal{X} \text{ covers } V\}.$$

⁴For $\mu = E(X)$ the Chernoff bound states $P(|X - \mu| \geq \delta\mu) \leq 2e^{-\delta^2\mu/3}$.

Proof. In [KvRW⁺05], the authors merely claim “a deterministic polynomial-time” algorithm without bounding the degree of the polynomial. We complete the proof by going through the individual steps of the algorithm and bounding the running time. Assume the input consists of $V = \{u_1, \dots, u_n\}$ and $\mathcal{X} = \{S_1, \dots, S_m\}$. In [KvRW⁺05, Section 5.1], the authors construct an equivalent ILP with $m + 1$ variables and $2n$ constraints whose matrix representation contains only entries from $\{-1, 0, 1\}$. As discussed in Section 6.1, the corresponding LP relaxation can be solved within time $\mathcal{O}((n + m)^{2.6}m)$ [Vai89]. This gives for each set S_i a fractional value $0 \leq x'_i \leq 1$. Then [KvRW⁺05, Section 5.2] defines numbers α and β and assigns each set S_i a probability $p_i := \min(1, \alpha x'_i)$. The paper then defines a pessimistic estimator

$$P(p_1, \dots, p_m) := 2 - \prod_{i=1}^n (1 - A_i) - \prod_{i=1}^n (1 - B_i)$$

where

$$A_i := \prod_{S_j \ni u_i} (1 - p_j) \quad \text{and} \quad B_i := \prod_{S_j \ni u_i} (1 + (\beta - 1)p_j).$$

The final derandomization procedure in [KvRW⁺05, Section 5.3] loops over $1 \leq i \leq m$ and rounds p_i to either 0 or 1 such that $P(p_1, \dots, p_m)$ is maximal. The evaluation of A_i and B_i takes time $\mathcal{O}(m)$, and thus each evaluation of $P(p_1, \dots, p_m)$ takes time $\mathcal{O}(nm)$. Since we loop over $1 \leq i \leq m$, the total running time of this step is $\mathcal{O}(nm^2)$. All in all, we arrive at running time of $\mathcal{O}((n + m)^{2.6}m)$. \square

Note that one may also directly apply the above pessimistic estimator of [KvRW⁺05] to derandomize Lemma 15. Instead, we prove Theorem 10 by a simple reduction to MMSC.

Proof of Theorem 10. Without loss of generality, we assume that $n \geq 5$. By Lemma 12 and Lemma 13, we can compute in time $\mathcal{O}(n^{9.8})$ a solution to the LP relaxation of cover-ILP(G, r, s) with value at most d^* . Next, Lemma 14 converts the solution in time $\mathcal{O}(n^2)$ into a fractional r -neighborhood cover (\mathcal{X}, f) with degree at most $d^* + 1$ and spread s . By Lemma 15, G has a weak r -neighborhood cover $\mathcal{Y} \subseteq \mathcal{X}$ with degree at most $36 \ln(n)(d^* + 1)$ and spread s . We will reduce the computation of a good approximation of this cover to the approximation of an MMSC solution.

We construct an MMSC instance (V', \mathcal{X}') where the universe $V' = V(G) \times \{0, 1\}$ contains for every vertex $v \in V(G)$ two copies $(v, 1)$ and $(v, 2)$, and the set system \mathcal{X}' contains for every cluster $X \in \mathcal{X}$ a set

$$X' = \{(v, 1) \mid v \in V(G), N_r[v] \subseteq X\} \cup \{(v, 2) \mid v \in X\}.$$

Claim 3. *For every $z \in \mathbb{N}$, $\mathcal{Y}' = \{X'_1, \dots, X'_\ell\} \subseteq \mathcal{X}'$ is a solution for the MMSC instance (V', \mathcal{X}') with $M(V', \mathcal{Y}') \leq z$ if and only if $\mathcal{Y} = \{X_1, \dots, X_\ell\} \subseteq \mathcal{X}$ is a weak r -neighborhood cover with degree z and spread s .*

Proof. Assume \mathcal{Y}' is an MMSC solution. For every vertex $v \in V(G)$, we have $(v, 1) \in X' \in \mathcal{Y}'$. By construction, we also have $N_r[v] \subseteq X \in \mathcal{X}$. Furthermore, for every $v \in V(G)$, we have that $(v, 2)$ is only contained in at most z sets of \mathcal{Y}' . Hence, v is contained in at most z clusters of \mathcal{Y} . As \mathcal{X} has spread s , the same holds for \mathcal{Y} . We conclude that \mathcal{X} is a weak r -neighborhood cover with degree z and spread s .

Assume \mathcal{Y} is a weak r -neighborhood cover with the degree z and spread s . Every $v \in V(G)$ appears in at most z clusters from \mathcal{Y} . Thus, $(v, 2)$ appears in at most z sets from \mathcal{Y}' . As $(v, 1) \in X'$ implies $(v, 2) \in X'$ for every $X' \in \mathcal{X}'$, the same holds for $(v, 1)$. For every $v \in V(G)$, its r -neighborhood is covered in some cluster $X \in \mathcal{Y}$. We therefore have $(v, 1) \in X' \in \mathcal{Y}'$ and also $(v, 2) \in X'$. We conclude that \mathcal{Y}' is an MMSC solution. ■

Having established the above reduction, we run Lemma 16 on the instance (V', \mathcal{X}') in time $\mathcal{O}((|V'| + |\mathcal{X}'|)^{2.6} |\mathcal{X}'|)$ and convert the output into a weak r -neighborhood cover $\mathcal{Y} \subseteq \mathcal{X}$ with spread s and degree at most $(1 + \mathcal{O}(1/\sqrt{z}))(\log(|V'|) + 1)z$, where $z = 36 \ln(n)(d^* + 1)$. The degree is bounded by $\mathcal{O}(\log(n)^2 + 1)d^*$. Since $|V'| = n$ and $|\mathcal{X}'| = \mathcal{O}(n^2)$, Lemma 16 runs in time $\mathcal{O}(n^{7.2})$, which dominates the time required to convert instances and solutions between the two problems. □

7 Weak Neighborhood Covers in Structurally Sparse Classes

In this section, we prove that structurally nowhere dense graph classes admit flip-closed sparse weak neighborhood covers.

Theorem 11. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon)$ such that for every $G \in \mathcal{C}$ there exists a weak r -neighborhood cover with degree at most $c(r, \varepsilon) \cdot |G|^\varepsilon$ and spread at most $34r$. In particular, \mathcal{C} admits flip-closed sparse weak neighborhood covers.*

The key to establishing our main theorem is a sparsification of the input graph by local contractions. To measure sparsity, we are going to use the weak coloring numbers $\text{wcol}_r(G)$, which were introduced by Kierstead and Yang [KY03]. We will formally define these numbers below. The weak coloring numbers are very useful for our purposes, since graphs with bounded weak coloring numbers admit neighborhood covers with good properties.

Lemma 17 ([GKS17, Lemma 6.10]). *Let $r \in \mathbb{N}$. Every graph G has an r -neighborhood cover with spread at most $2r$ and degree at most $\text{wcol}_{2r}(G)$.*

The key to proving Theorem 11 is the following.

Theorem 12. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $G \in \mathcal{C}$ there exists an 8-contraction $\text{contract}(G)$ of G , which is sparse in the following sense: for every $\varepsilon > 0$ and $r \in \mathbb{N}$ there exists $c(r, \varepsilon)$ such that for every $G \in \mathcal{C}$*

$$\text{wcol}_r(\text{contract}(G)) \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

We will first define local contractions and prove that we can lift sparse neighborhood covers from a contracted graph to the original graph. Then, by combining Lemma 17 and Theorem 12 we conclude Theorem 11.

7.1 Local Contractions and Weak Neighborhood Covers

The following definition of local contractions is the key to our approach.

Definition 6 (Contractions). Let G be a graph and let $A \subseteq V(G)$ and $\bar{A} = V(G) \setminus A$. The contraction $G[A]$ is defined as the graph obtained by contracting A into a new vertex v_A that is connected to every $w \in \bar{A}$ with $N[w] \cap A \neq \emptyset$. For disjoint subsets $A_1, \dots, A_l \subseteq V(G)$ such that each A_i is contained in a k -neighborhood of G , we call $G[A_1] \dots [A_l]$ a k -contraction of G .

Note that k -contractions are not necessarily depth- k minors, as the sets A_i may have large radius or may not even be connected (while they embed into low radius subgraphs). In particular, k -contractions in graphs from a nowhere dense class may not preserve nowhere denseness. However, we observe that a cover for a k -contraction of G can be lifted to a cover for G .

Lemma 18. *Let G' be a k -contraction of a graph G . If there exists a weak r -neighborhood cover with spread s and degree d for G' , then there exists a weak r -neighborhood cover with spread at most $(2k + 1) \cdot s$ and degree d for G .*

Proof. Let \mathcal{A} be the partition of the vertex set of G , such that each set $A \in \mathcal{A}$ is fully contained in a k -neighborhood in G and G' is obtained from G by contracting the parts of \mathcal{A} , that is, we have $V(G') = \mathcal{A}$ and $\{A_1, A_2\} \in E(G')$ if and only if there exist $v_1 \in A_1$ and $v_2 \in A_2$ such that $\{v_1, v_2\} \in E(G)$.

Let \mathcal{X}' be a weak neighborhood cover of spread s for G' . For a vertex $v \in V(G)$, denote by $\mathcal{A}(v)$ the set $A \in \mathcal{A}$ containing v . We construct the neighborhood cover \mathcal{X} for G as follows. For each cluster $X' \in \mathcal{X}'$, we define the cluster $X = \bigcup_{A \in X'} \{v : \mathcal{A}(v) = A\}$ and let $\mathcal{X} = \{X : X' \in \mathcal{X}'\}$.

We prove that \mathcal{X} is a weak r -neighborhood cover of G with spread at most $(2k + 1) \cdot s$ and the same degree as \mathcal{X}' .

Covering neighborhoods. We first show that for every vertex $w \in V(G)$ there exists a cluster $X \in \mathcal{X}$ containing $N_r^G[w]$. Let $X' = \text{cluster}'(\mathcal{A}(w)) \in \mathcal{X}'$ be the cluster containing $N_r^{G'}[\mathcal{A}(w)]$ and let X be the cluster from which X' was created. Let z be a vertex at distance at most r from w , which is witnessed by a path P in G between w and z of length at most r . Let $P' = \{\mathcal{A}(x) : x \in V(P)\}$. Since P is a path of length at most r , the subgraph induced by P' in G' is connected and has diameter at most r in G' . It follows that $\mathcal{A}(z)$ is at distance at most r from $\mathcal{A}(w)$ in G' and therefore contained in X' . By construction, z is contained in X . Thus, X covers $N_r^G[w]$.

Bounding spread. Let us now show that every cluster $X \in \mathcal{X}$ is a subset of a $((2k + 1) \cdot s)$ -neighborhood in G . Let $A_1 = \text{center}'(X')$, where X' is the cluster from which X was created. Choose an arbitrary vertex $v \in A_1$ and let $w \in X$. We show that w is at distance at most $(2k + 1) \cdot s$ from v in G . By construction, we have $\mathcal{A}(w) \in X'$. Since \mathcal{X}' has spread s , we have that A_1 and $\mathcal{A}(w)$ are at distance at most s in G' , witnessed by a path $P = (A_1 = \mathcal{A}(v), A_2, A_3, \dots, A_i = \mathcal{A}(w))$ of length $i \leq s$. Since any two vertices of A_j , $1 \leq j \leq i$, have distance at most $2k$ in G and there exist $v_j \in A_j$ and $v_{j+1} \in A_{j+1}$ with $\{v_j, v_{j+1}\} \in E(G)$, $1 \leq j < i$, we conclude that v and w have distance at most $(2k + 1) \cdot i$ in G .

Bounding degree. A vertex $v \in V(G)$ is contained in exactly those clusters X of \mathcal{X} such that $\mathcal{A}(v)$ is contained in X' of \mathcal{X}' . Since $\mathcal{A}(v)$ appears in at most d clusters of \mathcal{X}' , v appears in at most d clusters of \mathcal{X} . \square

7.2 Background on Structurally Sparse Graphs

The key to proving Theorem 12 is based on a structural characterization of structurally nowhere dense graph classes in terms of *quasi-bushes* [DGK⁺22a]. Let us recall the necessary background about nowhere dense and structurally nowhere dense graph classes.

7.2.1 Generalized Coloring Numbers and Sparsity Measures

A graph H is an r -shallow minor of a graph G if there is a set of pairwise disjoint vertex subsets $\{V_u \subseteq V(G)\}_{u \in V(H)}$ each of radius at most r such that if $\{u, v\} \in E(H)$, then there exists a vertex of V_u connected to a vertex of V_v . A graph H is an r -shallow topological minor if there is a set of vertices $\{p(u)\}_{u \in V(H)}$ and a set $\{P(u, v)\}_{\{u, v\} \in E(H)}$ of internally vertex disjoint paths of length at most $2r + 1$ such that $P(u, v)$ has endpoints $p(u)$ and $p(v)$.

In the following, $\nabla_r(G)$ and $\tilde{\nabla}_r(G)$ denote the maximal ratio of edges divided by vertices among all r -shallow minors and r -shallow topological minors in G , respectively.

We are going to use the weak coloring numbers, which were introduced by Kierstead and Yang [KY03]. Fix a graph G and an order \prec on the vertices of G . We say that a vertex u is *weakly r -reachable* from a vertex v if there is a path of length at most r between v and u such that u is the smallest vertex on the path (with respect to \prec). The set of all vertices weakly r -reachable from v is denoted by $\text{WReach}_r[G, \prec, v]$. We let

$$\text{wcol}_r(G, \prec) = \max_{v \in V(G)} |\text{WReach}_r[G, \prec, v]|,$$

and define the *weak r -coloring number* of G as

$$\text{wcol}_r(G) = \min_{\text{order } \prec \text{ on } V(G)} \text{wcol}_r(G, \prec).$$

By Lemma 17 every graph G has an r -neighborhood cover with spread at most $2r$ and degree at most $\text{wcol}_{2r}(G)$.

Additionally, $\text{adm}_r(G)$ stands for the r -admissibility of G , a sparsity measure we only use as an auxiliary concept to bound the other sparsity measures and whose exact definition does not concern us. The relations between these measures are nicely collected in chapter one and two of the sparsity lecture notes of Marcin and Michał Pilipeczuk [PPS20].

Proposition 1. *For every $r \in \mathbb{N}$ and graph G we have*

$$\text{wcol}_r(G) \leq 1 + r(\text{adm}_r(G) - 1)^{r^2} \quad [\text{PPS20, Corollary 2.7}] \quad (15)$$

$$\text{adm}_r(G) \leq 1 + 6r(\lceil \tilde{\nabla}_{r-1}(G) \rceil)^3 \quad [\text{PPS20, Lemma 3.2}] \quad (16)$$

$$\tilde{\nabla}_r(G) \leq \nabla_r(G) \quad \text{by definition} \quad (17)$$

$$\nabla_r(G) \leq \text{wcol}_{4r+1}(G) \quad [\text{PPS20, Lemma 3.1}] \quad (18)$$

Note that by Proposition 1 the values in $\text{wcol}_r(G)$, $\tilde{\nabla}_r(G)$ and $\nabla_r(G)$ are polynomially related (where the degree of the polynomial depends on r). This allows us to routinely use one notion to bound the other. The following link between average and minimum degree will also be useful.

Lemma 19 (Folklore, see for example (3.4) of [NDM12]). *Every graph with average degree at least d contains a subgraph with minimum degree at least $\frac{d}{2}$.*

7.2.2 Structurally Nowhere Dense Graph Classes

Nowhere dense graph classes can be defined using the weak r -coloring numbers.

Definition 7 (Nowhere Dense). A graph class \mathcal{C} is *nowhere dense* if for every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon)$ such that for every subgraph G of a graph from \mathcal{C} satisfies $\text{wcol}_r(G) \leq c(r, \varepsilon) \cdot |G|^\varepsilon$.

Structurally nowhere dense graph classes were defined in [GKN⁺20] as first-order transductions of nowhere dense classes. We will not rely on this definition, but instead use a decomposition result involving so-called almost nowhere dense quasi-bushes [DGK⁺22a] that we present below.

7.2.3 Almost Nowhere Dense Graph Classes

We can use the weak r -coloring numbers not only to define nowhere dense graph classes, but also to define the slightly more general notion of almost nowhere dense graph classes [DGK⁺22a].

Definition 8 (Almost Nowhere Dense). A graph class \mathcal{C} is *almost nowhere dense* if for every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon)$ such that every $G \in \mathcal{C}$ satisfies $\text{wcol}_r(G) \leq c(r, \varepsilon) \cdot |G|^\varepsilon$.

Observe that a hereditary graph class is almost nowhere dense if and only if it is nowhere dense. However, graphs from an almost nowhere dense class can contain dense subgraphs, for example, cliques of size $\log(n)$, if these appear as subgraphs of sufficiently large sparse graphs. By the previous Lemma 17, almost nowhere dense classes admit sparse neighborhood covers.

7.2.4 Quasi-Bushes

Following the local separator-based approach of [Dre21], quasi-bushes were introduced in [DGK⁺22a] to derive structural properties such as low shrubdepth covers for structurally nowhere dense graph classes. Let T be a rooted tree. For nodes $v, w \in V(T)$, we say that v is above w if it lies on the unique path from w to the root. Node v is below w if w is above v . Note that each node is above and below itself. For a rooted tree T and node w in the tree, the *ancestors* and *descendants* of w are all nodes above and below w in T (including w), respectively. Let us stress that each node is both an ancestor and descendant of itself. We write $v \leq_T w$ if v is an ancestor of w and $v <_T w$ if additionally $v \neq w$, in which case we call v a *strict ancestor* of w . Hence, the root is the smallest node of T and the leaves are the maximal nodes of T with respect to the tree order \leq_T . We define \geq_T and $>_T$ analogously as expected. For a vertex $w \in V(T)$, let $T(w)$ be the subtree of T rooted at w containing all descendants of w , and $L(w)$ be the set of leaves of T that are descendants of w .

Definition 9. A quasi-bush consists of

- a rooted tree T , represented by a directed graph in which all edges are directed away from the root,
- a set D of directed arcs (called *pointers*) from the leaves of T to the nodes of T (we require that every leaf points to the root of T)⁵,
- a labeling function $\lambda: D \rightarrow \{0, 1\}$.

A quasi-bush B defines a directed graph $G(B)$ whose vertices are the leaves of T and where the arc set is defined as follows: let u, v be two distinct leaves, and let w be the lowest (largest with respect to \leq_T) ancestor of v such that $(u, w) \in D$. Then (u, v) is an arc in $G(B)$ if and only if $\lambda((u, w)) = 1$. In this case w is called the *connection point* of (u, v) . Note that the connection point of an arc (u, v) is uniquely determined. Equivalently, (u, v) is an arc in $G(B)$ if on the shortest directed path from u to v in B that uses as its first arc a pointer $(u, w) \in D$ and all other arcs from T , the pointer (u, w) is labeled 1. We call this shortest path the *tunnel* of (u, v) in B . The *depth* of a quasi-bush is the depth of the tree T .

We say u has a *pointer* to v if $(u, v) \in D$. If $\lambda((u, v)) = 1$, we call the pointer *positive*. In this work $G(B)$ will always be treated as a directed graph. We will encode only undirected graphs, hence, $G(B)$ will be a symmetric, directed graph, that is, it will contain an arc (u, v) if and only if it contains the arc (v, u) . We will then make use of the fact that each arc (u, v) is encoded redundantly: via an ancestor of u and via an ancestor of v .

⁵In [DGK⁺22a] pointers can only point to internal nodes of T , for us it will be convenient to deal with the slightly more general definition and allow pointers to all vertices of T .

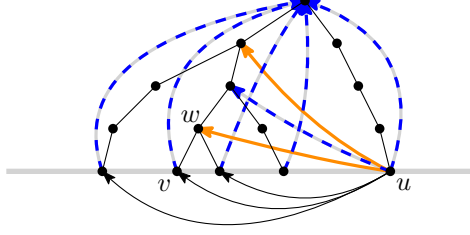


Figure 4: A quasi-bush (above the gray line) and the graph it describes (below the gray line). Dotted blue arcs represent pointers labeled 0 and orange arcs represent pointers labeled 1. The node w is the connection point of (u, v) .

The *Gaifman graph* of a quasi-bush is the (undirected) graph whose edge set consists of the tree-edges and the pointers. We extend all graph theoretic concepts to quasi-bushes via their Gaifman graphs. For example, the generalized coloring number $\text{wcol}_r(B)$ of a quasi-bush B refers to $\text{wcol}_r(H)$, where H is the Gaifman graph of G . We can thus speak, for example, about almost nowhere dense classes of quasi-bushes.

7.3 Quasi-Bush Decompositions of Structurally Nowhere Dense Classes

The central observation of [DGK⁺22a] shows that structurally nowhere dense classes can be described by almost nowhere dense quasi-bushes of bounded depth.

Theorem 13 ([DGK⁺22a, Theorem 3]). *Let \mathcal{D} be a structurally nowhere dense class of graphs. Then for every $G \in \mathcal{D}$, there is a quasi-bush B_G representing G such that the class of quasi-bushes $\{B_G : G \in \mathcal{D}\}$ has bounded depth and is almost nowhere dense.*

Since $\{B_G : G \in \mathcal{D}\}$ is almost nowhere dense, we know by definition that for every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon)$ such that for all $G \in \mathcal{D}$ there exists an ordering \prec such that $\text{wcol}_r(G, \prec) \leq c(r, \varepsilon) \cdot |G|^\varepsilon$. For our purposes, we additionally need that \prec can be chosen independently of r and ε and respecting the tree order, meaning that for all $u, v \in V(T)$ with $u \leq_T v$ we have $u \preceq v$.

Definition 10. Let $B = (T, D, \lambda)$ be a quasi-bush. We call an order \prec on $V(T)$ *ancestor respecting* if for all nodes $u, v \in V(T)$ with $u \leq_T v$, we have $u \preceq v$.

To see that we may make this assumption, we analyze the proofs of the full version [DGK⁺22b] of the paper that introduced quasi-bushes. The paper defines the notion of *r-separator quasi-bushes* in Definition 29. Given a graph $G \in \mathcal{D}$ derived from a graph H from a nowhere dense class \mathcal{C} via a transduction with Gaifman radius r , the paper then shows in the PROOF OF THEOREM 28 USING THEOREM 30, how to derive a quasi-bush of G (as we defined it above) via an *r-separator quasi-bush* of H with the same underlying tree and pointer set. Given an order \prec , an explicit construction of an *r-separator quasi-bush* of H , called $B_r^{\text{sep}}(H, \prec)$, is given in Definition 37. Lemma 39 then chooses the order \prec such that $\{B_r^{\text{sep}}(H, \prec) \mid H \in \mathcal{C}\}$ is almost nowhere dense. The (internal) nodes of $B_r^{\text{sep}}(H, \prec)$ are subsets of $V(H)$ with the property that for all nodes X, Y where X is an ancestor of Y we have $X \subseteq Y$. Lemma 39 first chooses a good weak-reachability order \prec on H and extends it to the internal nodes of $B_r^{\text{sep}}(H, \prec)$ so that for any two internal nodes $X, Y \subseteq V(H)$, $X \prec Y$ if and only if the maximum of X is smaller than the maximum of Y (with respect to \prec). This means, the order \prec places every ancestor before its descendant in $B_r^{\text{sep}}(H, \prec)$ and thus also in the quasi-bush of G . Furthermore, the construction is independent of r and ε , as desired. This yields the following result.

Theorem 14. Let \mathcal{D} be a structurally nowhere dense class of graphs. For every $G \in \mathcal{D}$, there exists a quasi-bush B_G representing G and an ancestor respecting order \prec_G such that $\{B_G : G \in \mathcal{D}\}$ has bounded depth. Additionally, for every $r \in \mathbb{N}$ and $\varepsilon > 0$, there exists $c(r, \varepsilon) \in \mathbb{N}$ such that for all $G \in \mathcal{D}$ we have

$$\text{wcol}_r(B_G, \prec_G) \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

Over the course of the following construction, we will modify quasi-bushes by contracting sets of nodes and adding additional pointers. To control the density of the resulting quasi-bushes, we need the following closure operation on Gaifman graphs of quasi-bushes.

Definition 11. Let $B = (T, D, \lambda)$ be a quasi bush rooted at vertex w_0 . We denote by B^* the undirected graph obtained by taking the Gaifman graph of B and adding to it for every pointer $(v, w) \in D$ from a leaf v to a vertex $w \in V(T)$, the edge set $\{\{v', w'\} \mid v' \leq_T v, w' \leq_T w\}$.

We also need a second kind of closure property that does not act on the Gaifman graph, but on the quasi-bush itself.

Definition 12. A quasi-bush (T, D, λ) is *upwards closed* if for all pointers $(u, w) \in D$ and ancestors $w' \leq_T w$ in T there also exists a pointer $(u, w') \in D$.

Note that $\lambda((u, w'))$ may differ from $\lambda((u, w))$.

Lemma 20. For every quasi-bush B of depth d with order \prec , there exists an upwards closed quasi-bush B' such that $G(B) = G(B')$, B and B' have the same underlying tree, and for every $r \in \mathbb{N}$, we have

$$\text{wcol}_r(B'^*, \prec) \leq \text{wcol}_{d \cdot r}(B, \prec).$$

In particular, since B and B' have the same underlying tree, if \prec is ancestor respecting on B , then the same is true on B' .

Proof. We construct B' from B adding for all $(u, w) \in D$ and ancestors w' of w in T , also the pointer (u, w') to D' . Remember that w is an ancestor of itself. We will specify the value of $\lambda'(u, w')$ soon. However, note that we already have – independent of λ' – that B' is upwards closed, and B and B' have the same underlying tree.

We define λ' such that $G(B) = G(B')$: We set $\lambda'(u, w') = \lambda(u, w'')$, where w'' is the lowest ancestor of w' with $(u, w'') \in D$. Since every node has a pointer to the root, such a node w'' always exists. Let u, v be two nodes. Let w' be the lowest ancestor of v with $(u, w') \in D'$. Let w'' be the lowest ancestor of w' with $(u, w'') \in D$ (note that $w' = w''$ is possible). Since $D \subseteq D'$, we observe that w'' is the lowest common ancestor of v with $(u, w'') \in D$. Thus, u and v are connected in $G(B)$ if and only if $\lambda((u, w'')) = 1$. By construction and choice of w' and w'' , $\lambda'((u, w')) = \lambda((u, w''))$, and thus u and v are connected in $G(B')$ if and only if $\lambda'((u, w')) = \lambda((u, w'')) = 1$. We conclude that u and v have the same connection in $G(B)$ as in $G(B')$.

Note that $B'^* = B^*$. It hence remains to show that for every $r \in \mathbb{N}$ we have $\text{wcol}_r(B^*, \prec) \leq \text{wcol}_{2d \cdot r}(B, \prec)$. Let $u, v \in V(T)$ be vertices such that v weakly r -reaches u in B^* with respect to \prec . This is witnessed by a path $P^* = (v, \dots, u)$ of length at most r in B^* where u is the smallest element in the path. We prove the claim by constructing a path P of length at most $2d \cdot r$ in B from v to u where u is the smallest element on the path according to \prec , witnessing that v weakly $(2d \cdot r)$ -reaches u in B .

We build P from P^* by replacing each edge of P^* with a path in B of length at most $2d$ whose vertices are all greater or equal to u with respect to \prec . Edges in P^* that are edges of T are also edges in B and do not need to be replaced. For every edge (a, b) in P^* that is not an edge in T , by

the definition of B^* , there exist descendants a' of a and b' of b in T that are connected by a pointer in B . Therefore there exists a path $P_{ab} = (a, \dots, a', b', \dots, b)$ in B that traverses T downwards from a to a' , uses the pointer to get to b' and traverses T upwards to get to b . Since P_{ab} only uses descendants of a and b and \prec is ancestor respecting, we know that all its vertices are greater or equal to u . Since T has depth at most d , P_{ab} has length at most $2d + 1$. We notice that the root vertex of T can be reached from every other vertex of T in at most d steps by traversing down T to a leaf and using a pointer to the root. We can therefore assume that neither a nor b is the root vertex, which improves our bound on the length of P_{ab} to $2(d - 1) + 1 \leq 2d$. \square

As a direct consequence of combining Theorem 14 and Lemma 20, we finally transform the decompositional result of [DGK⁺22a] into the shape we need.

Theorem 15. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $G \in \mathcal{C}$ there exists an upwards closed quasi-bush B_G representing G and an ancestor respecting order \prec_G such that $\{B_G : G \in \mathcal{C}\}$ has bounded depth. Additionally, for every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon) \in \mathbb{N}$ such that for all $G \in \mathcal{C}$ we have*

$$\text{wcol}_r(B_G^*, \prec_G) \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

7.4 Sibling Contractions in Quasi-Bushes

Our goal, as stated in Theorem 12, is to find for a graph G from a structurally nowhere dense graph class a sparse 8-contraction. We will construct this 8-contraction as a sequence of so-called *sibling contractions* on the quasi-bush of G , defined as follows. In the following, when we want to stress that we refer to the leaves $L(w)$ of a node w in the tree T of a specific bush $B = (T, D, \lambda)$ we write $L^B(w)$ for $L(w)$.

Let $B = (T, D, \lambda)$ be a quasi-bush. We call a set of vertices $S = \{a_1, \dots, a_k\} \subseteq V(T)$ with the same parent p in T a *sibling set*. Let $T(S)$ denote the forest consisting of the disjoint union of $T(a_1), \dots, T(a_k)$ and let $L^B(S) := L^B(a_1) \cup \dots \cup L^B(a_k)$. Let a be a vertex not in $V(T) \setminus V(T(S))$. We define the *sibling contraction* $B[a \leftarrow S] = (T', D', \lambda')$ as follows.

1. T' is built by removing the vertices $V(T(S))$ from T and inserting a below p as a leaf.
2. For the unmodified nodes, the pointers are inherited from B . To be more precise, for all nodes $v, w \in V(T) \cap V(T')$ we set $(v, w) \in D'$ if and only if $(v, w) \in D$ and $\lambda'((v, w)) = \lambda((v, w))$.
3. It remains to define the pointers from and to the new leaf a . For every $w \in V(T') \setminus \{a\}$ we define the following rules, that we will later refer to as (3a), (3b), (3c), (3d).
 - (a) We set $(a, w) \in D'$ if and only if $(u, w) \in D$ for some $u \in L^B(S)$.
 - (b) We set $(v, a) \in D'$ if and only if $(v, w) \in D$ for some $w \in V(T(S))$.
 - (c) We set $\lambda'((a, w)) = 1$ if and only if there exists $u \in L^B(S)$ and $v \in V(T) \setminus V(T(S))$ with $(u, v) \in E(G(B))$ such that w lies on the tunnel of (u, v) .
 - (d) We set $\lambda'((v, a)) = 1$ if and only if there is $u \in L^B(S)$ with $(v, u) \in E(G(B))$.

In case the name of the new vertex a is not important, we also write $B[S]$ instead of $B[a \leftarrow S]$. An example of a sibling contraction is given in Figure 5. Let us first argue that the result of applying a sibling contraction to a quasi-bush is still a quasi-bush representing an undirected graph, namely the undirected graph obtained from $G(B)$ by contracting the vertices of $L^B(S)$ into a single vertex.

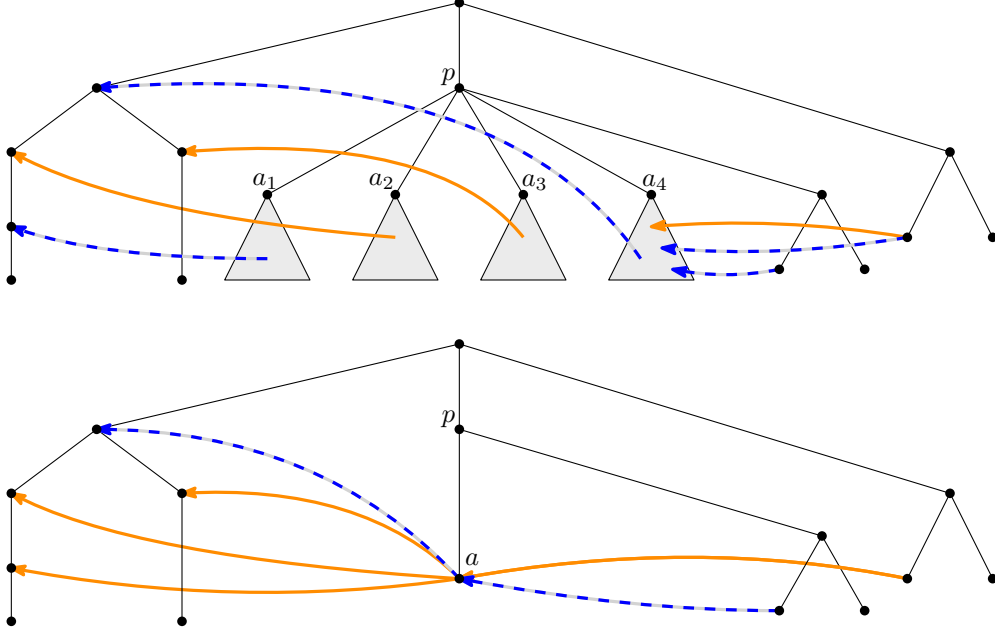


Figure 5: Quasi-bush B at the top and sibling contraction $B[a \leftarrow \{a_1, a_2, a_3, a_4\}]$ at the bottom.

Lemma 21. *Given a quasi-bush B representing an undirected graph and a sibling set S in B , then also $B[S]$ is a quasi-bush representing an undirected graph and we have*

$$G(B[S]) = G(B)[L^B(S)].$$

Proof. Let $B = (T, D, \lambda)$, $G = G(B)$, $B' = (T', D', \lambda') = B[a \leftarrow S]$, $G' = G(B')$, $T(S)$ and $L^B(S)$ be as above. Denote by p the parent of the vertices in S . By definition $V(G') = V(G[L^B(S)]) = (V(G) \setminus L^B(S)) \cup \{a\}$. Since a is a leaf, for any two vertices $u \neq a, v \neq a$, the bush induced by u and v and their ancestors is the same in B and B' and we have that u and v are connected (in both directions) in G' if and only if they are connected in G if and only if they are connected in $G[L^B(S)]$. It remains to verify the connections from and to a .

Claim 4. *Let $v \in V(G[L^B(S)]) = V(G')$ be a vertex adjacent to a in $G[L^B(S)]$. Then v is adjacent to a in G' , that is, $(a, v) \in E(G')$ and $(v, a) \in E(G')$.*

Proof. We first show $(a, v) \in E(G')$. By definition of $G[L^B(S)]$ there exists $u \in L^B(S)$ adjacent to v in G . Let w' be the connection point in B of (u, v) and let w be the lowest ancestor of v in T such that $(a, w) \in D'$. By (3a) such w exists and satisfies $w' \leq_T w \leq_T v$, with possibly $w = w'$. Then w lies on the tunnel of (u, v) and $\lambda'((a, w))$ is set to 1 by (3c). Since w is the lowest ancestor of v in T with $(a, w) \in D'$ we have $(a, v) \in E(G')$.

Now we show $(v, a) \in E(G')$. First assume $(v, a) \notin D'$. Then, by (3b), there is no $w \in V(T(S))$ such that $(v, w) \in D$. Since $(v, u) \in E(G)$, the connection point w of (v, u) satisfies $w \leq_T p$. This means w is the lowest ancestor of p in T with $(v, w) \in D$ and $\lambda((v, w)) = 1$. As B and B' agree on $V(T) \cap V(T')$, w also is the lowest ancestor of p in T' with $(v, w) \in D'$ and $\lambda'((v, w)) = 1$. Since p is the parent of a in T' and $(v, a) \notin D'$, w is the connection point of (v, a) and we have $(v, a) \in E(G')$.

Now assume $(v, a) \in D'$. We have that $\lambda'((v, a))$ is set to 1 as $(v, u) \in E(G)$ by (3d). It follows immediately that $(v, a) \in E(G')$. ■

Claim 5. *Let $v \in V(G[L^B(S)]) = V(G')$ be a vertex non-adjacent to a in $G[L^B(S)]$. Then v is non-adjacent to a in G' , that is, $(a, v) \notin E(G')$ and $(v, a) \notin E(G')$.*

Proof. By definition of $G[L^B(S)]$, v is non-adjacent to all of $L^B(S)$ in G .

We first show $(a, v) \notin E(G')$. Let w be the lowest ancestor of v in T' such that $(a, w) \in D'$. Then by (3a), w is the lowest ancestor of v in T such that $(u, w) \in D$ for some $u \in L^B(S)$. For the graph G , this means every $u \in L^B(S)$ is connected uniformly to all descendants of w : for all $u \in L^B(S)$ and descendants v_1, v_2 of w , $(u, v_1) \in E(G)$ if and only if $(u, v_2) \in E(G)$. Assume for contradiction that $(a, v) \in E(G')$ and therefore $\lambda'((a, w)) = 1$. By (3c) there exists leaves $u \in L^B(S)$ and $v' \in V(T) \setminus V(T(S))$ with $(u, v') \in E(G)$ such that w lies on the tunnel of (u, v') . Since v' and v are both descendants of w , also $(u, v) \in E(G)$. This contradicts $u \in L^B(S)$ and we conclude that $(a, v) \notin E(G')$.

We now show that $(v, a) \notin E(G')$. If $(v, a) \in D'$, then by (3d) we have $\lambda'((v, a)) = 0$, as there is no $u \in L^B(S)$ with $(v, u) \in E(G)$. Thus assume $(v, a) \notin D'$. Let w be the lowest ancestor of a such that $(v, w) \in D'$. By assumption $w \leq_T p$, hence also $(v, w) \in D$. Since v is non-adjacent to all of $L^B(S)$ in G we have $\lambda'((v, w)) = \lambda((v, w)) = 0$. As w is the lowest ancestor of a with $(v, w) \in D'$, we conclude that $(v, a) \notin E(G')$. ■

Now $G(B[S]) = G(B)[L^B(S)]$ follows. □

Note that sibling contractions preserve upwards closure, as shown in the next lemma.

Lemma 22. *Let B be an upwards closed quasi-bush and S be a sibling set. Then $B[a \leftarrow S]$ is upwards closed.*

Proof. It suffices to analyze the new pointers in $B[S]$ which point from and to the contracted vertex a . If a points to a vertex w in $B[S]$, then by (3a) some $u \in L^B(S)$ must also point to w in B . As B is upwards closed, u also points to all ancestors of w . It follows that also a points to all ancestors of w .

Conversely, if a vertex v points to a in $B[S]$, then by (3b) it must have pointed to a vertex u in a subtree beneath one of the siblings from S in B . As B is upwards closed it points to all ancestors of u in B , which is a superset of the ancestors of a in $B[S]$. Those pointers are still present in $B[S]$. □

7.5 Outline of the Contraction Procedure

Given a structurally nowhere dense graph class \mathcal{C} , we use Theorem 15 to associate with every $G \in \mathcal{C}$ an upwards closed bounded-depth quasi-bush B_G representing G . Furthermore, by that theorem, each B_G is equipped with an ancestor respecting order \prec_G witnessing at the same time that B_G^* is uniformly sparse, that is, for every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon) \in \mathbb{N}$ such that for all $G \in \mathcal{C}$ we have

$$\text{wcol}_r(B_G^*, \prec_G) \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

We will now perform a sequence of sibling contractions to arrive at a sparse 8-contraction of $G(B)$ whose existence we claim in Theorem 12. The sibling contractions will be performed in two phases, followed by an analysis.

- In phase one, we will simplify the quasi-bush by eliminating certain patterns called *coat hangers*. This phase only uses sibling contractions with sibling sets of size one. The goal of this phase is to gain ultimate control over tunnels.

Recall that for every arc (u, v) of $G(B)$ the shortest directed path from u to v in B that uses as its first arc a pointer $(u, w) \in D$ and all other arcs from T is the tunnel of (u, v) in B , where w is called the connection point of (u, v) . After eliminating coat hangers, we will have the property that every tunnel for (u, v) will either be of the form u, v , that is, the endpoint v of (u, v) is the connection point of the arc itself, of the form u, w, v , that is, the connection point w of (u, v) is the parent of v .

- Hence, after the first phase, for all arcs (u, v) of $G(B)$ with a connection point w (possibly equal to v) we have u as an in-neighbor (in D) of w and v as an out-neighbor (in T) of w (or $v = w$). This naturally motivates the definition of sets $\text{In}(w)$ and $\text{Out}(w)$.

In phase two, we contract sibling sets of unbounded size to bound the cardinality of the sets $\text{In}(w)$ and $\text{Out}(w)$.

- Afterwards, we argue that a sparse quasi-bush with small sets $\text{In}(w)$ and $\text{Out}(w)$ for all necessarily represents a sparse graph, thereby proving Theorem 12.

The sibling contractions insert new nodes and edges into the bush. For our final analysis, we require that this does not make the resulting quasi-bush too dense. In phase one, we start with a quasi-bush $B = B_0$ and repeatedly contract sibling sets of size one. This means we contract quasi-bushes $B_i = (T_i, D_i \lambda_i)$ into quasi-bushes $B_{i+1} = (T_{i+1}, D_{i+1} \lambda_{i+1}) = B_i \llbracket a \leftarrow \{a\} \rrbracket$ for some node a . To bound the density of these operations, we note that all newly introduced edges in B_{i+1} originate from edges in B_i with endpoints below a , and thus the closure B_{i+1}^* is a subgraph of the closure B_i^* . Since by Theorem 15 $\text{wcol}_r(B_0^*)$ is small, $\text{wcol}_r(B_i^*), \text{wcol}_r(B_{i+1}^*), \dots$ are small, too.

However, this is not enough. To start phase two, we not only require that the weak coloring numbers are small, but even that they are small with respect to an *ancestor respecting* order. For directed graphs H' and H , let us write $H' \sqsubset H$ if $V(H') \subseteq V(H)$ and $E(H') \subseteq E(H)$. Note that this means nothing else than that H' is a subgraph of H , but we want to stress that $V(H') \subseteq V(H)$ and we want to distinguish between isomorphic subgraphs.

We will show that $B_{i+1}^* \sqsubset B_i^*$ and $T_{i+1} \sqsubset T_i$. Note that the latter means T_{i+1} is a subtree of T_i . Theorem 15 gives us an ancestor respecting order \prec such that $\text{wcol}_r(B_0^*, \prec)$ is small. Since $B_{i+1}^* \sqsubset B_i^*$, the coloring numbers with respect to \prec stay small, and since $T_{i+1} \sqsubset T_i$, the order \prec stays ancestor respecting. Thus, the following lemma is crucial to prepare the quasi-bush for the second phase.

Lemma 23. *For every quasi-bush B , $a \in V(T)$ and $B' = (T', D', \lambda') := B \llbracket a \leftarrow \{a\} \rrbracket$ we have $B'^* \sqsubset B^*$ and $T' \sqsubset T$.*

Proof. It is clear from the construction of B' that $T' \sqsubset T$ and that $V(B'^*) = V(B') \subseteq V(B) = V(B^*)$. Recall that B^* is a supergraph of the Gaifman graph of B and therefore contains undirected edges. To also show that $E(B'^*) \subseteq E(B^*)$, consider an edge $\{u, v\} \in E(B'^*)$. If $\{u, v\}$ originates from an arc of T' then, with $T' \sqsubset T$, the edge $\{u, v\}$ is also in $E(B^*)$. Therefore, assume $\{u, v\} \in E(B'^*)$ to not originate from a tree-arc.

Otherwise, the edge $\{u, v\}$ originates from a pointer (in any direction) between u_0 and v_0 in B' with $u \leq_{T'} u_0$ and $v \leq_{T'} v_0$. If neither u_0 nor v_0 is equal to the leaf a , then the pointer between u_0 and v_0 is also present in B . Since $u \leq_T u_0$ and $v \leq_T v_0$, B^* also contains the edge $\{u, v\}$.

Otherwise, assume by symmetry that $u_0 = a$. During the construction of B' , by (3a) and (3b), there must have a pointer between some $a_0 \in T(a)$ and v_0 . Since $u \leq_T u_0 = a \leq_T a_0$ and $v \leq_T v_0$, B^* also contains the edge $\{u, v\}$. \square

Unfortunately, $B'^* \sqsubset B^*$ only holds for sibling contractions $B' = B[a \leftarrow S]$ with $|S| = 1$ in general. In phase two, we will contract multiple sibling sets of larger cardinality. We will still ensure that our operations preserves the sparsity of our quasi-bushes, by

1. only contracting non-overlapping sibling sets, which we will call *independent*, and
2. embedding the contracted quasi-bush into a slightly larger but still sparse graph, which we will call *copy product*.

Definition 13 (Copy product). Given a graph G , we define its *copy product* $\text{cp}(G)$ as the graph obtained by taking the disjoint union of G with its copy G' and connecting every node v from G with its copy $\text{cp}(v)$ in G' and making v and $\text{cp}(v)$ true twins. If \prec is an order on G , then $\text{cp}(\prec)$ is the order on $\text{cp}(G)$ obtained from \prec by inserting each element $\text{cp}(v)$ right before v .

Observe that the copy product is equal to the lexicographic product of G with K_2 . Taking the lexicographic product with a clique of bounded size is known to preserve sparsity (see for example [NDM12, Proposition 4.6.]).

Lemma 24. *For every graph G , every order \prec and every $r \in \mathbb{N}$ we have*

$$\text{wcol}_r(\text{cp}(G), \text{cp}(\prec)) \leq 2 \cdot \text{wcol}_r(G, \prec).$$

Proof. By definition, if $\text{WReach}_r[G, \prec, v] = \{v_1, \dots, v_l\}$ then both $\text{WReach}_r[\text{cp}(G), \text{cp}(\prec), v]$ and $\text{WReach}_r[\text{cp}(G), \text{cp}(\prec), \text{cp}(v)]$ are subsets of $\{v_1, \dots, v_l, \text{cp}(v_1), \dots, \text{cp}(v_l)\}$. \square

Definition 14. Two sibling sets S_1 and S_2 are *independent* in a quasi-bush B if they have different parent vertices and $L^B(a)$ and $L^B(b)$ are disjoint for all $a \in S_1$ and $b \in S_2$ (see Figure 6). Let S_1, \dots, S_l be pairwise independent sibling sets in B and let $\{a_1, \dots, a_l\}$ be distinct nodes not in B . It follows that the quasi-bushes $B_0 = B$ and $B_i = B_{i-1}[a_i \leftarrow S_i] = B[a_1 \leftarrow S_1] \dots [a_i \leftarrow S_i]$ for $i \in [l]$ are valid sibling contractions: a_i is not contained in B_{i-1} and S_i is a sibling set in B_{i-1} for every $i \in [l]$. We write $B[a_1 \leftarrow S_1, \dots, a_l \leftarrow S_l]$ to denote iterated contraction of pairwise independent sibling sets.

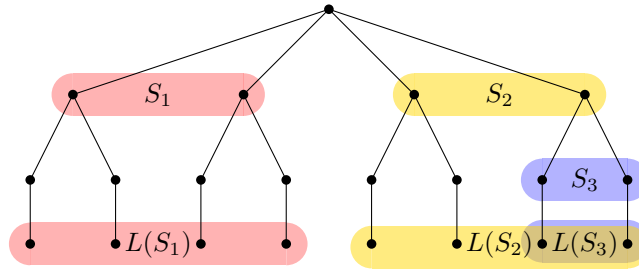


Figure 6: Three sibling sets S_1, S_2, S_3 in a quasi-bush (pointers omitted). S_1 and S_2 are not independent because they share the same parent. S_2 and S_3 are not independent because their leaf sets $L(S_2)$ and $L(S_3)$ overlap. S_1 and S_3 are independent.

In phase two, we will only contract pairwise independent sibling sets and use the following observation to bound the sparsity of the contraction.

Lemma 25. *For every quasi-bush B_0 , pairwise independent sibling sets S_1, \dots, S_l with parents p_1, \dots, p_l , and sibling contraction $B_l := B[\text{cp}(p_1) \leftarrow S_1, \dots, \text{cp}(p_l) \leftarrow S_l]$ we have $B_l^* \sqsubset \text{cp}(B_0^*)$.*

Proof. Let $B_0 = (T_0, D_0, \lambda_0) = B$ and $B_i = (T_i, D_i, \lambda_i) = B_{i-1} \llbracket \text{cp}(p_i) \leftarrow S_i \rrbracket$ for all $i \in [l]$.

Observation 2. For all $i \in [l]$ we have $V(T_i) \setminus \{\text{cp}(p_1), \dots, \text{cp}(p_i)\} \subseteq V(T_0)$.

Then it is clear that $V(B_l^*) = V(T_l) \subseteq V(T_0) \cup \{\text{cp}(p_1), \dots, \text{cp}(p_l)\} \subseteq V(\text{cp}(B_0^*))$. We now prove by induction for all $i \in [l]$ that every edge in B_i^* is also an edge of $\text{cp}(B^*)$. For $i = 0$ we have $B_0 = B$ and $B_0^* = B^* \sqsubset \text{cp}(B^*)$ holds by definition.

Assume the statement holds for B_i^* . Let $\{u, v\}$ be an edge in B_{i+1}^* but not in B_i^* , as otherwise the statement follows by induction. If $\{u, v\}$ originates from a tree arc in B_{i+1}^* , then we have that the edge $\{u, v\} = \{p_i, \text{cp}(p_i)\}$ is contained in $\text{cp}(B^*)$ by definition of the copy product. Otherwise, $\{u, v\}$ originates from a pointer (in any direction) between nodes $u_0 \geq_{T_{i+1}} u$ and $v_0 \geq_{T_{i+1}} v$ in B_{i+1} .

Claim 6. Either $u_0 = \text{cp}(p_{i+1})$ or $v_0 = \text{cp}(p_{i+1})$.

Proof. Assume towards a contradiction that neither u_0 nor v_0 is $\text{cp}(p_{i+1})$. Then also neither u nor v is $\text{cp}(p_{i+1})$ and both are contained in B_i . As the sibling contraction only modified pointers from and to $\text{cp}(p_{i+1})$, the pointer between u_0 and v_0 is present in B_i already. Therefore, $\{u, v\}$ is an edge in B_i^* , contradicting our previous assumption. ■

By symmetry, we can assume that $u_0 = \text{cp}(p_{i+1})$. As there is a pointer between $u_0 = \text{cp}(p_{i+1})$ and v_0 and by (3a) and (3b), there must exist a vertex $u'_0 \geq_{T_i} p_{i+1}$ and a pointer between u'_0 and v_0 in B_i . By $p_{i+1} \leq_{T_i} u'_0$ and $v \leq_{T_i} v_0$, the edges $\{p_{i+1}, v_0\}$ and $\{p_{i+1}, v\}$ exist in B_i^* .

Claim 7. $u = u_0 = \text{cp}(p_{i+1})$.

Proof. Assume towards a contradiction that $u \neq \text{cp}(p_{i+1}) = u_0$. Then $u \leq_{T_{i+1}} \text{cp}(p_{i+1})$. As p_i is the parent of $\text{cp}(p_{i+1})$ in T_{i+1} we conclude that $u \leq_{T_{i+1}} p_{i+1}$. Since T_{i+1} differs from T_i only in the subtree below p_i , we also have $u \leq_{T_i} p_{i+1}$. Together with the edge $\{p_{i+1}, v\}$ in B_i^* , this implies the edge $\{u, v\}$ in B_i^* , contradicting our previous assumption. ■

Recall that $u'_0 \geq_{T_i} p_{i+1}$. As the sibling sets S_1, \dots, S_{i+1} are independent, we must have that $u'_0 \notin \{\text{cp}(p_1), \dots, \text{cp}(p_i)\}$ and $u'_0 \in V(T_0)$, by Observation 2.

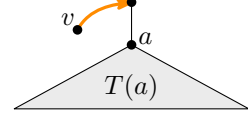
Assume $v_0 \notin \{\text{cp}(p_1), \dots, \text{cp}(p_i)\}$. Again, we know that $v_0 \in V(T_0)$, by Observation 2. The pointer between u'_0 and v_0 is therefore already present in B . Then B_0^* contains the edge $\{p_{i+1}, v\}$. By definition of the copy product, $\text{cp}(B_0^*)$ contains the edge $\{u = \text{cp}(p_{i+1}), v\}$ as desired.

Finally, assume $v_0 = \text{cp}(p_j)$ for some $j \in [i]$. Then the pointer between v_0 and u'_0 in B_i must have been introduced in B_j . By (3a) and (3b), there must be a pointer between some $v'_0 \geq_{T_{j-1}} p_j$ and u'_0 in B_{j-1} . As the sibling sets S_1, \dots, S_j are independent, we have that $v'_0 \notin \{\text{cp}(p_1), \dots, \text{cp}(p_{j-1})\}$ and $v'_0 \in V(T_0)$ by Observation 2. Then the pointer between u'_0 and v'_0 is also present in B_0 . It follows that the edge $\{p_{i+1}, p_j\}$ is in B_0^* . Recall that $v \leq_{T_{i+1}} v_0$. If $v = v_0 = \text{cp}(p_j)$, then the edge $\{\text{cp}(p_{i+1}), \text{cp}(p_j)\} = \{u, v\}$ in $\text{cp}(B_0^*)$ results from the edge $\{p_{i+1}, p_j\}$ in B_0^* and the definition of the copy product. Otherwise, $v \leq_{T_{i+1}} \text{cp}(p_j)$ and since $\text{cp}(p_j)$ is inserted below p_j , also $v \leq_{T_{i+1}} p_j$. As the nodes on the path from p_j to the root of T_0 remain unchanged during the sibling contraction, we also have that $v \in T_0$ and $v \leq_{T_0} p_j$. From the edge $\{p_{i+1}, p_j\}$ in B_0^* we conclude that the edge $\{p_{i+1}, v\}$ exists in B_0^* , and by the definition of the copy product, $\text{cp}(B_0^*)$ contains the edge $\{u = \text{cp}(p_{i+1}), v\}$ as desired. □

7.6 Phase One: Eliminating Coat Hangers

Definition 15. Given a quasi-bush $B = (T, D, \lambda)$, a *coat hanger* is a subtree $T(a)$ rooted at a node a , such that there exists a leaf v (not necessarily contained in $T(a)$) with the following properties

- a is not the root and v positively points to the parent of a ,
- v does not point to any node from $T(a)$,
- a is a non-leaf node.



We call a coat hanger $T(a)$ *maximal*, if there exists no coat hanger $T(a')$ such that a' is an ancestor of a in T .

Observation 3. For every quasi-bush $B = (T, D, \lambda)$ and maximal coat hangers $T(a_1)$ and $T(a_2)$ with $a_1 \neq a_2$, we have that $V(T(a_1)) \cap V(T(a_2)) = \emptyset$.

We will show below that quasi-bushes without coat hangers have a nice structure. In order to eliminate coat hangers, we start by showing that sibling contractions do not introduce new coat hangers.

Lemma 26. Let $B = (T, D, \lambda)$ be an upwards closed bush, S be a sibling set, and $B' = (T', D', \lambda') := B \llbracket S \rrbracket$. If $T'(a')$ is a coat hanger in B' for some $a' \in V(T')$, then also $a' \in V(T)$ and there exists $a \leq_T a'$ such that $T(a)$ is a coat hanger in B .

Proof. As $T'(a')$ is a coat hanger in B' , a' is a non-leaf in B' and there exists a leaf v' with a positive pointer to the ancestor w' of a' , and no pointer to any node in $T'(a')$. Since both a' and w' are non-leaves in T' , they were not contracted in the sibling contraction and were therefore present in B already. The node v' was either already present in B , or was introduced by the contraction.

First, assume that v' was already present in B . Then the pointer $\lambda'((v', w')) = 1$ was inherited from B . Since v' has no D' -pointer to any node in $T'(a')$, we conclude by (3b) that v' also has no D -pointer to any node in $T(a')$. Hence, $T(a')$ is a coat hanger in B .

Now assume that v' was introduced by the contraction and the D' -pointer (v', w') was newly introduced. Since $\lambda'((v', w')) = 1$, we conclude by (3c) that w' lies on a tunnel, witnessed by a leaf $v \in T(S)$ and a node $w \leq_T w'$ in B such that

- (1) v has a positive D -pointer to w , and
- (2) there is no D -pointer from v to any node a with $w <_T a \leq_T w'$.

Let $a \leq_T a'$ be the non-leaf child of w on the path from w to a' . Let us show that v has no pointer to any node from $T(a)$ in B .

If $w \neq w'$, then $w <_T w'$. As a is a child of w , it satisfies $w <_T a \leq_T w'$ and then v has no pointer to a in B by (2). As B is upwards closed, v then also cannot have a pointer to a node from $T(a)$ in B .

On the other hand, if $w = w'$, then $a = a'$ and a D -pointer from v to $T(a) = T(a')$ would, as B is upwards closed, imply a D -pointer from v to a' in B and by (3a) also a D' -pointer from v' to a' in B' . A contradiction to v' not pointing to $T'(a')$.

We conclude that in B , v has a positive D -pointer to w , which has a non-leaf child a such that v points to no element from $T(a)$. By definition, $T(a)$ is a coat hanger in B . \square

We derive the following simple corollary of Lemma 26.

Corollary 1. *If an upwards closed quasi-bush contains no coat hangers, then neither does any of its sibling contractions.*

We now prove the central observation of the first contraction phase, which constructs a 1-contraction that removes all coat hangers without making the quasi-bush too dense.

Lemma 27. *For every upwards closed quasi-bush $B = (T, D, \lambda)$, there exists an upwards closed coat hanger free quasi-bush $B' = (T', D', \lambda')$ such that $G(B')$ is a 1-contraction of $G(B)$, $B'^* \sqsubset B^*$ and $T' \sqsubset T$.*

Proof. Let $T(a_1), \dots, T(a_l)$ be the maximal coat hangers of B . We set $B_0 = (T, D, \lambda) := B$ and for every $i \in [l]$ we define $B_i = (T_i, D_i, \lambda_i) := B_{i-1} \llbracket a_i \leftarrow \{a_i\} \rrbracket$. We set $B' := B_l$. Lemma 23 gives us $B'^* = B_l^* \sqsubset B_{l-1}^* \sqsubset \dots \sqsubset B_0^* = B^*$. Then, by transitivity of the subgraph relation $B'^* \sqsubset B^*$, it similarly follows by Lemma 23 that $T' \sqsubset T$. By Lemma 22, B' remains upwards closed. It remains to argue that B' is coat hanger-free and that $G(B')$ is a 1-contraction of $G(B)$.

To prove that B' is coat hanger-free, we show by induction on $i \in [l]$ that

$$\text{for every coat hanger } T(b) \text{ in } B_i \text{ we have } a_j \leq_{T_i} b \text{ for some } i < j \leq l.$$

The statement trivially holds for $B_0 = B$. Assume it holds for B_i and let us prove it for $B_{i+1} = B_i \llbracket a_{i+1} \leftarrow \{a_{i+1}\} \rrbracket$. To this end, let $T(b')$ be a coat hanger in B_{i+1} . By Lemma 26, $b' \in V(T_i)$ and there exists $b \leq_{T_i} b'$ such that $T(b)$ is a coat hanger in B . By induction, $a_j \leq_{T_i} b$ for some $i < j \leq l$ and thus also $a_j \leq_{T_i} b'$. Since $a_j, b' \in V(T_{i+1}) \cap V(T_i)$ and sibling contractions preserve the tree of preserved nodes, it follows that $a_j \leq_{T_{i+1}} b'$ for some $i < j \leq l$. Notice that a_{i+1} is a leaf in B_{i+1} , while, since $T(b')$ is a coat hanger, b is not a leaf in B_{i+1} . Hence, we know that $a_{i+1} \not\leq_{T_{i+1}} b'$. Therefore, $a_j \leq_{T_{i+1}} b'$ for some $i+1 < j \leq l$, which proves the statement.

It remains to show that $G(B')$ is a 1-contraction of $G(B)$. Let $L^{B_i}(a)$ be the leaves below a in B_i . By repeated application of Lemma 21, we have

$$G(B') = G(B) \llbracket L^{B_1}(a_1) \rrbracket \dots \llbracket L^{B_l}(a_l) \rrbracket.$$

It is easy to see that $L^{B_i}(a_i) = L^B(a_i)$. The sets $L^B(a_1), \dots, L^B(a_l)$ are by Observation 3 pairwise disjoint. Since $T(a_i)$ is a coat hanger in B , each set $L^B(a_i)$ furthermore induces a subgraph of radius one in $G(B)$. According to Definition 6, $G(B') = G(B) \llbracket L^B(a_1) \rrbracket \dots \llbracket L^B(a_l) \rrbracket$ is a 1-contraction of $G(B)$. \square

It is now time to take a step back and to combine Theorem 15, and Lemma 27 into the following statement that finalizes the coat hanger elimination of phase one.

Lemma 28. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $G \in \mathcal{C}$, there exists a coat hanger free upwards closed quasi-bush B_G representing a 1-contraction of G and an ancestor respecting order \prec_G such that $\{B_G : G \in \mathcal{C}\}$ has bounded depth. Additionally, for every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon) \in \mathbb{N}$ such that for all $G \in \mathcal{C}$ we have*

$$\text{wcol}_r(B_G^*, \prec_G) \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

Proof. Theorem 15 gives for every $G \in \mathcal{C}$ an upwards closed quasi-bush B_G representing G and an ancestor respecting order \prec_G . For every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon) \in \mathbb{N}$ such that for all $G \in \mathcal{C}$ we have $\text{wcol}_r(B_G^*, \prec_G) \leq c(r, \varepsilon) \cdot |G|^\varepsilon$.

By Lemma 27, we can obtain from B_G an upwards closed coat hanger free quasi-bush B'_G representing a 1-contraction of G . Since $B'^*_G \sqsubset B_G^*$, we still have for every $r \in \mathbb{N}$ and $\varepsilon > 0$ that $\text{wcol}_r(B'^*_G, \prec_G) \leq c(r, \varepsilon) \cdot |G|^\varepsilon$. Since furthermore $T' \sqsubset T$, the order \prec_G remains ancestor respecting in B'_G . \square

7.7 Phase Two: Bounding In and Out Sets

One notices that representing a dense graph using a sparse quasi-bush requires “central hubs” that act as a connection point of many tunnels (see Figure 7). The following definition of In- and Out-sets measure how central a node w is.

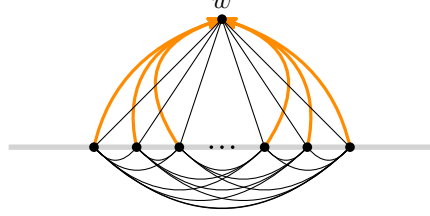


Figure 7: A sparse quasi-bush representing a clique. Here the node w acts as central hub through which many tunnels pass.

Definition 16. For every quasi-bush $B = (T, D, \lambda)$ and node $w \in V(T)$, we define

$$\begin{aligned} \text{In}(w, B) &:= \{u \mid w \text{ is connection point of some arc } (u, v) \in E(G(B))\}, \\ \text{Out}(w, B) &:= \{v \mid w \text{ is connection point of some arc } (u, v) \in E(G(B))\}. \end{aligned}$$

Note that in general, the nodes in $\text{In}(w, B)$ all have a positive pointer to w , while the nodes in $\text{Out}(w, B)$ are descendants of w . The main reason why we eliminated coat hangers in the previous phase is the following observation guaranteeing that also the nodes in $\text{Out}(w, B)$ have distance at most one from w in the bush.

Lemma 29. *Let $B = (T, D, \lambda)$ be an upwards closed quasi-bush containing no coat hangers and let $w \in V(T)$. Every node $v \in \text{Out}(w, B)$, $v \neq w$ is a child of w .*

Proof. Let $v \in \text{Out}(w, B)$. Then w is the connection point of some edge $(u, v) \in E(G(B))$. Assume towards a contradiction that w is not equal to v and not the parent of v . By definition of a connection point, u does not point to another node on the path from w to v in T . Let w' be the child of w on this path. As v is neither equal to nor a child of w' , we have $w <_T w' <_T v$. The node u does not point to w' and as B is upward closed, u does not point to any node from $T(w')$, as otherwise w would not be a connection point of (u, v) . Since $w' <_T v$, we conclude that w' is a non-leaf and hence $T(w')$ is a coat hanger. A contradiction. \square

Intuitively speaking, we will prove in the next subsection that a sparse quasi-bush B can only represent a dense graph if there are “central hubs” w for which both the set $\text{In}(w, B)$ and the set $\text{Out}(w, B)$ are of polynomial size. The following lemma eliminates these hubs, guaranteeing that B represents a sparse graph.

Lemma 30. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $G \in \mathcal{C}$ there exists a coat hanger free upwards closed quasi-bush B' representing an 8-contraction of G with the following properties. For every $r \in \mathbb{N}$ and every $\varepsilon > 0$ there exist $c(r, \varepsilon)$ and $t(\varepsilon)$ such that for every $G \in \mathcal{C}$ and every $r \in \mathbb{N}$ we have*

$$\text{wcol}_r(B') \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

Additionally, for every node w in B' we have

$$\text{In}(w, B') \leq t(\varepsilon) \cdot |G|^\varepsilon \quad \text{or} \quad \text{Out}(w, B') \leq t(\varepsilon) \cdot |G|^\varepsilon.$$

Proof. Let $G \in \mathcal{C}$. We first show how to construct the corresponding quasi-bush B' and then prove its properties.

Construction. We first apply Lemma 28 to obtain a coat hanger free upwards closed quasi-bush $B = (T, D, \lambda)$ representing a 1-contraction of G together with an ancestor respecting order \prec of $V(T)$ such that for all $r \in \mathbb{N}$ and $\varepsilon > 0$ we have $\text{wcol}_r(B^*, \prec) \leq c'(r, \varepsilon) \cdot |G|^\varepsilon$.

We iteratively process the non-leaf nodes of T in the order given by \prec and modify the quasi-bush B in every step as follows. Let $B_i = (T_i, D_i, \lambda_i)$ be the quasi-bush right before the i th processing step and let $W_i = \{w_1, \dots, w_{i-1}\}$ be set of nodes processed so far. We start with $B_1 = B$ and $W_1 = \emptyset$.

In the i th processing step, we choose w_i as the smallest (with respect to \prec) non-leaf node of T_i that is contained in $V(T)$ and larger (with respect to \prec) than all nodes of W_i . Note that w_1 is the root of T . If no such w_i exists we finish the processing by setting $B' := B_i$ and $W' := W_i$. Otherwise, we set $W_{i+1} := W_i \cup \{w_i\}$ and describe in the following how to derive B_{i+1} :

Choose the sibling set $S_i \subseteq V(T_i)$ to be a maximal (but possibly empty) set of children of w_i in T_i such that there exists some neighborhood of radius 2 in $G(B)$ that contains all leaves $L^{B_i}(S_i)$ below S_i in T_i . If S_i is empty, we simply set $B_{i+1} := B_i$. Otherwise, we perform a sibling contraction and set $B_{i+1} := B_i \llbracket \text{cp}(w_i) \leftarrow S_i \rrbracket$ and continue with the next processing step. Note that this is a valid sibling contraction since $\text{cp}(w_i)$ is not contained in B_i . This completes the description of the construction of B' .

Claims. Let k be the number of processing steps performed on B such that $W' = \{w_1, \dots, w_k\}$. We control the above construction using the following statements.

Claim 8. *For all $i \in [k]$ and $v \in V(T_i) \cap V(T)$ with $w_i \preceq v$ we have that $T_i(v) = T(v)$. In particular, v is a leaf in T_i if and only if it is a leaf in T .*

Proof. Assume towards a contradiction that $w_i \preceq v$ and $T_i(v)$ differs from $T(v)$. Since the sibling contractions only modify the subtree of the nodes that is currently processed, $T(v)$ must have been altered during the processing of a node $w_j \prec w_i \preceq v$, which, because \prec is ancestor respecting, must be an ancestor of v in T . During the sibling contraction below w_j , every child contained in S_j is deleted together with its subtree, while the subtrees of children not contained in S_j are left untouched. The existence of v in T_i then proves that the subtree containing v was not modified. A contradiction. ■

Claim 9. *For all $i \in [k]$ and $u, v \in V(T_i) \cap V(T)$ we have $(u, v) \in D_i$ if and only if $(u, v) \in D$. Additionally, we have $\lambda_i(u, v) = \lambda(u, v)$.*

Proof. The only pointers that are modified by the sibling contractions point towards and away from newly inserted nodes, which are not contained in $V(T)$. The pointers between nodes from $V(T_i) \cap V(T)$ therefore remain unmodified. ■

Claim 10. *The sets S_1, \dots, S_k are pairwise independent in B . Hence,*

$$B' = B \llbracket \text{cp}(w_1) \leftarrow S_1, \dots, \text{cp}(w_k) \leftarrow S_k \rrbracket.$$

Proof. Each set S_i contains children of w_i in B_i . By Claim 8 we have $T_i(w_i) = T(w_i)$. Hence, w_i is not only the parent of S_i in T_i , but also in T .

Fix $i < j \in [k]$ and let us show that S_i and S_j are independent in B . Clearly, S_i and S_j have different parent nodes $w_i \prec w_j$ in T . Let $a_i \in S_i$ and $a_j \in S_j$. Since the subtree $T_i(a_i) = T(a_i)$ was removed during the processing of w_i , it has no overlap with $T_j(a_j) = T(a_j)$, which was processed later. This implies $L^B(a_i) \cap L^B(a_j) = \emptyset$. By Definition 14, S_i and S_j are independent in B . ■

Claim 11. B_i is upwards closed and contains no coat hangers for all $i \in [k]$. In particular, $\text{Out}(w, B_i)$ contains only children of w .

Proof. B_i is obtained from B via a sequence of sibling contractions. As B originates from Lemma 28, B is upwards closed and contains no coat hangers. For each contraction, Lemma 22 preserves upwards closure and Corollary 1 preserves coat hanger freeness. ■

Claim 12. W' is the set of non-leaf nodes of B' . Furthermore, W' is contained in the set of non-leaves of B_i for every $i \in [k]$.

Proof. Let $w_i \in W'$ be the non-leaf of B_i chosen at the beginning of the i th processing step. During the i th processing step, either some children of w_i were merged into a leaf, or B_i was not changed. Either way, w_i remains a non-leaf in B_{i+1} . In the following processing steps only nodes that appear after w_i in \prec are processed. Since \prec is ancestor respecting, in particular, no ancestor of w_i is processed and w_i must have survived as a non-leaf in B' .

Let w be a non-leaf in B' . During the processing steps that transform B into B' , all nodes that are newly inserted are leaves, and since only sibling sets below non-leaf are contracted, these nodes stay leaves. Therefore, w exists as a non-leaf in B_i for every $i \in [k]$. In particular, w appears in $B_1 = B$ and is therefore ordered by \prec . If w is the root node of T , then w is trivially contained in W' . Hence, assume w is not the root node of T . Then there exists a node $w_{i-1} \in W'$ that is maximal with $w_{i-1} \prec w$. In the i th processing step, we try to pick w_i as the smallest (with respect to \prec) non-leaf node of T_i that is contained in $V(T)$ and larger (with respect to \prec) than w_{i-1} , or terminate if no such w_i exists. Note that w is a candidate for w_i and thus w_i exists and $w_i \preceq w$. If $w_i \prec w$ we get a contradiction to our choice of w_{i-1} . We therefore have $w = w_i \in W'$. ■

Claim 13. For every $w \in W'$ and $i \in [k-1]$, we have $|\text{In}(w, B_i)| \geq |\text{In}(w, B_{i+1})|$.

Proof. By Claim 12, w is a non-leaf in both B_i and B_{i+1} . We have that $B_{i+1} = B_i \llbracket S_i \rrbracket$. If $B_{i+1} = B_i$ there is nothing to show. Otherwise, there exists a new leaf $\text{cp}(w_i)$ below w_i in B_{i+1} . Let $u \in \text{In}(w, B_{i+1})$. By Definition 16 and since w is a non-leaf, there exists a leaf $v \neq w$ with $v \in \text{Out}(w, B_{i+1})$. By Claim 11 and Lemma 29, v is a child of w in B_{i+1} . In summary, u has a positive pointer to w and no pointer to a leaf-child u of w .

First assume $u \neq \text{cp}(w_i)$. If u has a pointer to every leaf below w in B_i , then the same must be true in B_{i+1} . This is a contradiction to our assumption that u has no pointer to v . Therefore, also in B_i there exists a leaf below w which v does not point to and we have $u \in \text{In}(w, B_i)$. Now assume $u = \text{cp}(w_i)$. Since u has a positive pointer to w , by definition of the sibling contraction, there must exist some contracted node $u' \in \text{In}(w, B_i)$ that is no longer present in B_{i+1} so we have $u' \notin \text{In}(w, B_{i+1})$. Combining both cases, we get $|\text{In}(w, B_i)| \geq |\text{In}(w, B_{i+1})|$. ■

We will now use these claims to prove the desired properties of B' . By Claim 11, B' is upwards closed and contains no coat hangers. The remaining three properties are shown in the remaining three paragraphs.

8-Contraction. We argue that $G(B')$ is an 8-contraction of G . By repeated application of Lemma 21, we have

$$G(B') = G(B) \llbracket L^{B_1}(S_1) \rrbracket \dots \llbracket L^{B_l}(S_l) \rrbracket.$$

We guaranteed during the construction that each set $L^{B_i}(S)$ is contained in a radius 2 neighborhood in $G(B)$. By Claim 8, $L^{B_i}(S_i) = L^B(S_i)$. The sets S_1, \dots, S_l are by Claim 10 independent sibling sets in B and thus by Definition 14, $L^B(S_1), \dots, L^B(S_l)$ are pairwise disjoint. According to Definition 6, $G(B') = G(B) \llbracket L^B(S_1) \rrbracket \dots \llbracket L^B(S_l) \rrbracket$ is a 2-contraction of $G(B)$. Since $G(B)$ itself is a 1-contraction of G , $G(B')$ is an 8-contraction of G .

Sparsity. We know that for every $r \in \mathbb{N}$ and $\varepsilon > 0$ that $\text{wcol}_r(B^*, \prec) \leq c'(r, \varepsilon) \cdot |G|^\varepsilon$, where $c'(r, \varepsilon)$ is the function originating from the invocation of Lemma 28 in the paragraph *Construction* at the beginning of the proof. We set $c(r, \varepsilon) = 2c'(r, \varepsilon)$. Then By Lemma 24, we have for the copy product $\text{cp}(B^*)$,

$$\text{wcol}_r(\text{cp}(B^*), \text{cp}(\prec)) \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

As shown in Claim 10, B' is obtained from B by a sibling contraction of pairwise independent sibling sets. By Lemma 25, the Gaifman graph of B' is a subgraph of $\text{cp}(B^*)$, and thus $\text{wcol}_r(B') \leq \text{wcol}_r(\text{cp}(B^*))$. This yields the desired bound on the weak coloring numbers of B' .

Size of In and Out Sets. In the following, remember that $W' = \{w_1, \dots, w_k\}$ is the set of processed nodes, and for $i \in [k]$, B_i is the quasi-bush right before processing w_i . As shown in Claim 10, we further have that $B' = B \llbracket S_1, \dots, S_k \rrbracket$ is a sibling contraction of pairwise independent sibling sets S_1, \dots, S_k in B . Choose an arbitrary $\varepsilon > 0$. Let $p := c(1, \varepsilon) \in \mathbb{N}$. As observed above,

$$\text{wcol}_1(\text{cp}(B^*), \text{cp}(\prec)) \leq p \cdot |G|^\varepsilon. \quad (19)$$

We want to show that for every node w of B' either

$$|\text{In}(w, B')| \leq 5p \cdot |G|^\varepsilon \quad \text{or} \quad |\text{Out}(w, B')| \leq 4p \cdot |G|^\varepsilon + 1.$$

We can then choose $t(\varepsilon)$ such that either $|\text{In}(w, B')| \leq t(\varepsilon) \cdot |G|^\varepsilon$ or $|\text{Out}(w, B')| \leq t(\varepsilon) \cdot |G|^\varepsilon$. If w is a leaf, then $|\text{Out}(w, B')| \leq 1$. Hence, assume w is a non-leaf and $|\text{In}(w, B')| > 5p \cdot |G|^\varepsilon$, as otherwise there is nothing to show.

By Claim 11, Lemma 29 and since w is a non-leaf, we get that $\text{Out}(w, B')$ is a subset of the children of w in B' . In order to bound $\text{Out}(w, B')$, it therefore suffices to show that w has at most $4p \cdot |G|^\varepsilon + 1$ children in B' . Since w is a non-leaf in B' , by Claim 12 we must have $w = w_i \in W'$ for some $i \in [k]$. Remember that B_i is the quasi-bush right before w_i was processed, and B_{i+1} is the quasi-bush after w_i was processed. It is easy to see that w_i has at least as many children in B_{i+1} as in B' .

It therefore suffices to show that w_i has at most $4p \cdot |G|^\varepsilon + 1$ children in B_{i+1} .

We now want to partition the nodes in $\text{In}(w_i, B_i)$, using the order \prec . However, \prec is only defined for nodes of B . We obtained $B_i = B \llbracket \text{cp}(w_1) \leftarrow S_1, \dots, \text{cp}(w_{i-1}) \leftarrow S_{i-1} \rrbracket$ from B by performing sibling contractions. For every $j < i$, B_i contains a leaf $\text{cp}(w_j)$, created during the contraction of S_j , which is not present in B and therefore not ordered by \prec . To circumvent this problem, we turn to the extended order $\text{cp}(\prec)$ given by Definition 13. We notice that all the nodes in B_i are also present in $\text{cp}(B)$. Therefore, $\text{cp}(\prec)$ completely orders the nodes of B_i , such that $V(T) \cap V(T_i)$ is ordered as in \prec and every node $\text{cp}(w_j) \in V(T) \setminus V(T_i)$ is the immediate predecessor of its parent w_j .

We know by (19) that, in terms of coloring numbers, $\text{cp}(\prec)$ is a good ordering for $\text{cp}(B^*)$. Lemma 25 states that $B_i^* \sqsubset \text{cp}(B^*)$ and thus the bounds given by (19) transfer to B_i , that is,

$$\text{wcol}_1(B_i, \text{cp}(\prec)) \leq p \cdot |G|^\varepsilon. \quad (20)$$

We can now partition the nodes in $\text{In}(w_i, B_i)$ into sets $\text{In}_{\prec}(w_i, B_i)$ and $\text{In}_{\succ}(w_i, B_i)$ depending on whether they are smaller or greater than w_i with respect to $\text{cp}(\prec)$. Let us first bound the size of $\text{In}_{\prec}(w_i, B_i)$. By Definition 16, every node in $\text{In}_{\prec}(w_i, B_i) \subseteq \text{In}(w_i, B_i)$ has a pointer to w in B_i . All nodes in $\text{In}_{\prec}(w_i, B_i)$ are smaller than w_i , and thus w_i weakly 1-reaches all of $\text{In}_{\prec}(w_i, B_i)$. It follows by (20) that

$$|\text{In}_{\prec}(w_i, B_i)| \leq \text{wcol}_1(B_i, \text{cp}(\prec)) \leq p \cdot |G|^\varepsilon.$$

We assumed $|\text{In}(w, B')| \geq 5p \cdot |G|^\varepsilon$. By Claim 13, $|\text{In}(w, B_i)| > |\text{In}(w, B')|$ and thus

$$|\text{In}_{\succ}(w_i, B_i)| = |\text{In}(w, B')| - |\text{In}_{\prec}(w, B')| > 4p \cdot |G|^\varepsilon. \quad (21)$$

We finally arrive at the central argument behind this proof. For this, we partition the children of w_i in T_i into a set A containing every child a such that less than half of the nodes in $\text{In}_{\succ}(w_i, B_i)$ have a pointer to a and a remaining set \bar{A} . See also Figure 8.

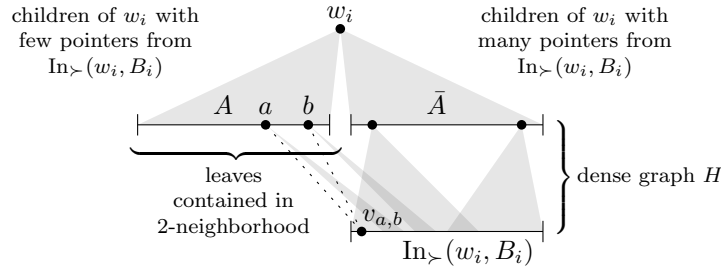


Figure 8: The central argument behind the proof of Lemma 30. The leaves below A are contained in a 2-neighborhood and hence can be contracted. Either the remaining children \bar{A} of w_i or the set $\text{In}_{\succ}(w_i, B_i)$ must be small, since otherwise a large dense graph H arises.

Let us argue that all the leaves below A are contained in a neighborhood of radius 2 in $G(B)$. By definition of A , for every pair $a, b \in A$, there must be $v_{a,b} \in \text{In}_{\succ}(w_i, B_i)$, neither pointing to a nor b . By upwards closure, it follows that $v_{a,b}$ points to no node from $T_i(a)$ and $T_i(b)$. Since $v_{a,b}$ has a positive pointer to w_i , we know that $v_{a,b}$ is connected in $G(B_i)$ to all the leaves $L^{B_i}(a)$ and $L^{B_i}(b)$ below a and b .

We argue that this is the case also in $G(B)$. In the order $\text{cp}(\prec)$, $v_{a,b}$ comes after w_i . As argued before, this means $v_{a,b} \in V(T)$ and $w_i \prec v_{a,b}$. Applying Claim 8 to $v_{a,b}$ gives us that $v_{a,b}$ is not only a leaf in T_i but also in T . Claim 8 also states that $T_i(w_i) = T(w_i)$, and thus in particular, $L^{B_i}(a) = L^B(a)$ and $L^{B_i}(b) = L^B(b)$. By Claim 9, the pointers in B between $T(w_i)$ and $v_{a,b}$ are the same as in B_i . It follows that also in $G(B)$, $v_{a,b}$ is connected to all of $L^{B_i}(a)$ and $L^{B_i}(b)$. Since this holds for every pair $a, b \in A$, the leaves below A (which are the same in T_i as in T) are contained in a neighborhood of radius 2 in $G(B)$.

Remember that $S_i \subseteq V(T_i)$ is chosen as a maximal (but possibly empty) set of children of w_i in T_i such that there exists some neighborhood of radius 2 in $G(B)$ that contains all leaves $L^{B_i}(S_i)$ below S_i in T_i . We just proved that A is a candidate for S_i and thus, when processing w_i , we have contracted a set S_i containing at least $|A|$ many of its children into a single new node $\text{cp}(w_i)$ below w_i . This means the number of children of w_i in B_{i+1} is at most $|\bar{A}| + 1$.

It therefore suffices to show that $|\bar{A}| \leq 4p \cdot |G|^\varepsilon$.

We bound the size of \bar{A} next. Assume towards contradiction $|\bar{A}| > 4p \cdot |G|^\varepsilon$. By (21), also $|\text{In}_{\succ}(w_i, B_i)| > 4p \cdot |G|^\varepsilon$. Consider the bipartite graph H defined by the pointers between \bar{A} in one part (called the upper part) and $\text{In}_{\succ}(w_i, B_i)$ in the other part (called the lower part). See also

Figure 8 for a depiction of H . Remember that by definition, every node in \bar{A} has an incoming pointer from at least half of the nodes in $\text{In}_{\succ}(w_i, B_i)$. Thus, every node from the upper part is connected to at least half of the nodes of the lower part. We bound the number of edges per vertex in H by

$$\frac{|\bar{A}| \cdot \frac{1}{2} |\text{In}_{\succ}(w_i, B_i)|}{|\bar{A}| + |\text{In}_{\succ}(w_i, B_i)|} \geq \frac{|\bar{A}| \cdot \frac{1}{2} |\text{In}_{\succ}(w_i, B_i)|}{2 \max(|\text{In}_{\succ}(w_i, B_i)|, |\bar{A}|)} = \frac{1}{4} \min(|\bar{A}|, |\text{In}_{\succ}(w_i, B_i)|) > p \cdot |G|^\varepsilon.$$

The graph H therefore has an average degree larger than $2p \cdot |G|^\varepsilon$ and by Lemma 19 a subgraph H' with minimum degree larger than $p \cdot |G|^\varepsilon$. This implies $\text{wcol}_1(B_i) \geq \text{wcol}_1(H') > p \cdot |G|^\varepsilon$, which is a contradiction to (20).

Thus, we must have $|\bar{A}| \leq 4p \cdot |G|^\varepsilon$ and can conclude that w_i has at most $4p \cdot |G|^\varepsilon + 1$ many children in B_{i+1} , which gives us the desired upper bound on $|\text{Out}(w_i, B')|$. \square

7.8 Sparsity of the Contraction

We will now show that the constructed 8-contraction is sparse. We first show that it does not contain large subdivided cliques. In the following, we will write $\text{In}(w)$ and $\text{Out}(w)$ instead of $\text{In}(w, B)$ and $\text{Out}(w, B)$, when the quasi-bush B will be clear from the context.

Lemma 31. *For every $r, m \in \mathbb{N}^+$ with $m \geq 20r$ and every upwards closed, coat hanger free quasi-bush B , if*

- $\text{wcol}_{9r}(B, \prec) \leq m$, and
- for every $w \in V(T)$ either $|\text{In}(w)| \leq m$ or $|\text{Out}(w)| \leq m$,

then $G(B)$ does not contain an r -shallow topological clique minor of size m^7 .

Proof. Assume towards a contradiction that $G(B)$ contains an r -shallow topological clique minor of size m^7 . It consists of m^7 principal vertices and $\binom{m^7}{2}$ pairwise vertex disjoint subdivision paths of length at most $(2r + 1)$ connecting all pairs of principal vertices. We orient each subdivision path $P_{uv} = (u, \dots, v)$ if $u \prec v$.

We assign each principal vertex u a node $p(u) \in V(T)$: If the parent w of u has $u \in \text{Out}(w)$ and $|\text{Out}(w)| \leq m$, we set $p(u) = w$; otherwise, we set $p(u) = u$. We set $A_u := \{u, p(u)\}$ and say that two distinct vertices u and v overlap if $A_u \cap A_v \neq \emptyset$. Let us count how many other principal vertices v may overlap with u . Since $u \neq v$, we know that the sets A_u and A_v intersect in a vertex $p(v) = p(u)$ different from u and v . By definition, $v \in \text{Out}(p(v)) = \text{Out}(p(u))$ and $|\text{Out}(p(u))| \leq m$ and thus every vertex u overlaps with at most m other vertices. We greedily pick from the m^7 principal vertices a set of pairwise non-overlapping principal vertices S of size m^5 . This is possible since $m \geq 10$ and thus

$$\frac{m^7}{m + 1} \geq m^5 = |S|.$$

For $(u, v) \in E(G(B))$ let $q(u, v)$ be the connection point of (u, v) . Since the bush is upwards closed and coat hanger-free, by Lemma 29, we have that $q(u, v)$ is either equal to v or the parent of v . As a direct consequence of Definition 16, we obtain the important observation that

$$u \in \text{In}(q(u, v)) \text{ and } v \in \text{Out}(q(u, v)). \quad (22)$$

We continue to work in the smaller subdivision spanned by the principal vertices from S . For a subdivision path $P = (v_1, \dots, v_l)$, we define $A_P := \{v_2, q(v_2, v_3), v_3, \dots, q(v_{l-2}, v_{l-1}), v_{l-1}\}$ to be the vertices on the tunnels connecting the internal nodes of P . Note that A_P is empty if and only if P has length one, that is, directly connects two principal vertices via an edge.

We say a principal vertex $u \in S$ and a subdivision path P *overlap* if $A_u \cap A_P \neq \emptyset$. Since u is no internal node of P , if v and P overlap then $u \neq p(u) = q(v', v)$ for some internal nodes v, v' of P . By (22) and the definition of $p(u)$, $v \in \text{Out}(q(v', v)) = \text{Out}(p(u))$ and $|\text{Out}(p(u))| \leq m$. Since v is an internal vertex of P and all subdivision paths are internally vertex disjoint, u overlaps with at most m subdivision paths. With $m \geq 10$, the set \mathcal{P}_1 of subdivision paths that run between two vertices from S , but overlap with no vertex from S therefore has size at least

$$|\mathcal{P}_1| \geq \binom{|S|}{2} - |S| \cdot m = \frac{m^5 \cdot (m^5 - 1)}{2} - m^5 \cdot m \geq m^9.$$

We say two subdivision paths P and P' *overlap*, if $A_P \cap A_{P'} \neq \emptyset$. Let us count how many other subdivision paths $P' = (v'_1, \dots, v'_l)$ may overlap with P . We know that P and P' are internally vertex disjoint, and thus the sets A_P and $A_{P'}$ intersect at a vertex $q = q(v'_{j-1}, v'_j)$. We use (22) and distinguish two cases:

- $|\text{Out}(q)| \leq m$. Since $v'_j \in \text{Out}(q)$ and all subdivision paths are internally vertex disjoint, there are at most m possible choices for P' such that A_P and $A_{P'}$ intersect in q .
- $|\text{In}(q)| \leq m$. Since $v'_{j-1} \in \text{In}(q)$ and all subdivision paths are internally vertex disjoint, again, there are at most m possible choices for P' such that A_P and $A_{P'}$ intersect in q .

Since there are at most $2r$ possible choices of q , in total, there can be at most $2rm$ other paths P' that overlap with P . It follows with $m \geq 20r$ that we can greedily pick a maximal subset $\mathcal{P}_2 \subseteq \mathcal{P}_1$ of pairwise non-overlapping paths of size at least

$$|\mathcal{P}_2| \geq \frac{|\mathcal{P}_1|}{2rm + 1} \geq \frac{m^9}{2rm + 1} \geq m^7.$$

If P is a subdivision path and $v \in P$ is a start- or endpoint of P we say that P is *incident* to v . Let v' be the neighbor of v in P . Since v is a start- or endpoint, v' is uniquely defined. Note that if P contains no inner vertices, that is, P has length one, then v' is the other start- or endpoint of the path. We say P *privately connects* to v , if $q(v', v) \in A_v$.

Remember that $q(v', v)$ is either equal to v or the parent w of v . If $q(v', v) \notin A_v$ then $q(v', v) = w$ and $p(v) = v$. This means by definition of $p(v)$ that $v \notin \text{Out}(w)$ or $|\text{Out}(w)| > m$. By (22), $v \in \text{Out}(q(v', v)) = \text{Out}(w)$ and thus $|\text{Out}(w)| > m$, which implies $|\text{In}(w)| \leq m$. Again by (22), $v' \in \text{In}(q(v', v)) = \text{In}(w)$. Note that the subdivision paths that are incident to v all differ in their neighbor v' of v . It follows that for every principal vertex $v \in S$, all but at most m of the subdivision paths incident to v privately connect to it. With $m \geq 10$, the subset $\mathcal{P}_3 \subseteq \mathcal{P}_2$ of subdivision paths that privately connect to both of their endpoints in S therefore has size at least

$$|\mathcal{P}_3| \geq m^7 - |S| \cdot m = m^7 - m^5 \cdot m > m^6.$$

Let us now take a look at the auxiliary graph H whose vertex set is S and where two vertices $u \prec v$ are connected if $P_{uv} \in \mathcal{P}_3$. We will argue that H is an $(r + 1)$ -shallow minor (but not necessarily a topological minor) of the Gaifman graph of B : note that the sets A_v and A_P are all pairwise disjoint for all $v \in S$ and $P \in \mathcal{P}_3$. Let $P = P_{uv} \in \mathcal{P}_3$. By our choice of \mathcal{P}_3 , if $A_P \neq \emptyset$ there are pointers from A_P to A_u and from A_P to A_v . If $A_{P_{uv}} = \emptyset$, then there are pointers between A_u

and A_v . Since A_P contains at most $2(r-1)$ vertices and A_u and A_v contain at most two vertices, each vertex in A_P has distance at most $r+1$ either to u or to v in the Gaifman graph of B . Thus, the sets A_v and A_P for $v \in S$ and $P \in \mathcal{P}_3$ together witness that H is an $(r+1)$ -shallow minor of the Gaifman graph of B . The density of H is

$$\frac{|\mathcal{P}_3|}{|S|} > \frac{m^6}{m^5} = m,$$

and thus $\nabla_{r+1}(B) > m$. By (18) from Proposition 1,

$$\text{wcol}_{9r}(B) \geq \text{wcol}_{4r+5}(B) = \text{wcol}_{4(r+1)+1}(B) \geq \nabla_{r+1}(B) > m.$$

This violates our assumption and finishes the proof of the lemma. \square

We will use the following tool that lets us build large topological clique minors in sufficiently dense graphs.

Lemma 32 ([Dvo07, Lemma 3.15]). *Let $\rho' \in \mathbb{N}$. There exist n'_0 and ρ'' such that all graphs G on $n \geq n'_0$ vertices with minimum degree at least $n^{1/\rho'}$ contain a ρ'' -shallow topological clique minor of size at least $n^{1/\rho''}$.*

We will also need the following transitivity observation about shallow topological minors.

Lemma 33 (See Proposition 4.2 of [NDM12]). *If a graph A contains a graph B as a b -shallow topological minor and B itself contains a graph C as a c -shallow topological minor, then A contains C as a $4bc$ -shallow topological minor.*

Using these two lemmas, we lift Lemma 31 towards thresholds that are polynomial in the number of leaves.

Lemma 34. *For every $\rho \in \mathbb{N}$ there exists $\mu, n_0 \in \mathbb{N}$ such that for every n -leaf upwards closed, coat hanger free quasi-bush B with $n \geq n_0$, if*

- $\text{wcol}_\mu(B) \leq n^{1/\mu}$, and
- for every $w \in V(T)$ either $\text{In}(w) \leq n^{1/\mu}$ or $\text{Out}(w) \leq n^{1/\mu}$,

then $\text{wcol}_\rho(G(B)) \leq n^{1/\rho}$.

Proof. We fix $\rho \in \mathbb{N}$. We will start with $n_0 = 1$ and increase it over the course of this proof as needed. The constant $\mu \in \mathbb{N}$ will be chosen later on. Let us pick a quasi-bush B with $n \geq n_0$ leaves satisfying the prerequisites of the lemma. Let $G = G(B)$ be the represented graph, which has n vertices. Assume towards a contradiction that $\text{wcol}_\rho(G) > n^{1/\rho}$. Combining the inequalities (15) and (16) of Proposition 1, yields

$$n^{1/\rho} < \text{wcol}_\rho(G) \leq 1 + \rho(6\rho)^{\rho^2} (\lceil \tilde{\nabla}_\rho(G) \rceil)^{3\rho^2}.$$

We can therefore choose ρ' and n_0 (as a function of ρ) such that with $n \geq n_0$, $\tilde{\nabla}_\rho(G) \geq n^{1/\rho'}$. In other words, G contains a ρ -shallow topological minor with at least $n^{1/\rho'}$ edges per vertex, and thus with average degree at least $2n^{1/\rho'}$. By Lemma 19, there exists a subgraph H of this ρ -shallow topological minor with minimum degree at least $n^{1/\rho'}$. We next want to apply Lemma 32 to H . Choose ρ'' and n'_0 as a function of ρ' according to this lemma. We update n_0 such that $n_0^{1/\rho'} \geq n'_0$. Since H has at least $n_0^{1/\rho'} \geq n'_0$ vertices, the prerequisites of Lemma 32 are met. Thus, H contains

a ρ'' -shallow topological clique minor of size at least $n^{1/\rho''}$. Since H itself is a ρ -shallow topological minor G and by transitivity of the shallow topological minor relation (Lemma 33), G contains an $4\rho\rho''$ -shallow topological clique minor of size $n^{1/\rho''}$.

We choose $r := 4\rho\rho''$, $\mu := 9r$ and $m := \lfloor n^{1/\mu} \rfloor$. Thus, G contains a r -shallow topological clique minor of size at least $n^{1/\rho''} \geq n^{1/r} > n^{7/9r} = n^{7/\mu} \geq m^7$. By our assumptions, for every $w \in V(T)$ either $\text{In}(w) = \lfloor \text{In}(w) \rfloor \leq \lfloor n^{1/\mu} \rfloor = m$ or $\text{Out}(w) = \lfloor \text{Out}(w) \rfloor \leq \lfloor n^{1/\mu} \rfloor = m$. We increment n_0 such that $m = \lfloor n^{1/9r} \rfloor \geq \lfloor n_0^{1/9r} \rfloor \geq 20r$. Since all the prerequisites of Lemma 31 are met, $\text{wcol}_{9r}(B) > m = \lfloor n^{1/\mu} \rfloor$. Since $\text{wcol}_{9r}(B)$ is an integer, also $\text{wcol}_{9r}(B) > n^{1/\mu}$. This contradiction to our assumption finishes the proof of the lemma. \square

7.9 Wrapping Up

We are ready to prove Theorem 12, which we restate for convenience.

Theorem 12. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $G \in \mathcal{C}$ there exists an 8-contraction $\text{contract}(G)$ of G , which is sparse in the following sense: for every $\varepsilon > 0$ and $r \in \mathbb{N}$ there exists $c(r, \varepsilon)$ such that for every $G \in \mathcal{C}$*

$$\text{wcol}_r(\text{contract}(G)) \leq c(r, \varepsilon) \cdot |G|^\varepsilon.$$

Proof. We first show that for every $\rho \in \mathbb{N}$ there exists $n_0(\rho)$ such that for all $G \in \mathcal{C}$ with $|G| \geq n_0(\rho)$ we have $\text{wcol}_\rho(G(B)) \leq |G|^{1/\rho}$, where B is the upwards closed, coat hanger-free quasi-bush representing an 8-contraction of G that we obtain via Lemma 30. Fix $\rho \in \mathbb{N}$. Let μ and n_0 (as a function of ρ) be the corresponding constants from Lemma 34. As we obtained B via Lemma 30,

- $\text{wcol}_\mu(B) \leq c(\mu, 1/2\mu) \cdot |G|^{1/2\mu}$,
- for every node w of B either $\text{In}(w) \leq t(1/2\mu) \cdot |G|^{1/2\mu}$ or $\text{Out}(w) \leq t(1/2\mu) \cdot |G|^{1/2\mu}$.

We can increase n_0 (as a function of $c(\mu, 1/2\mu)$ and $t(1/2\mu)$) such that for $|G| \geq n_0$

- $\text{wcol}_\mu(B) \leq n^{1/\mu}$,
- for every node w of B either $\text{In}(w) \leq n^{1/\mu}$ or $\text{Out}(w) \leq n^{1/\mu}$.

Since B is upwards closed and contains no coat hangers, we have by Lemma 34 for $|G| \geq n_0$,

$$\text{wcol}_\rho(G(B)) \leq |G|^{1/\rho}.$$

Now for $r \in \mathbb{N}$ and $\varepsilon > 0$ let $\rho = \max\{r, \varepsilon^{-1}\}$ and $c(r, \varepsilon) = n_0(\rho)$. Then for all graphs $G \in \mathcal{C}$ with $|G| \geq n_0(\rho)$,

$$\text{wcol}_r(G(B)) \leq \text{wcol}_\rho(G(B)) \leq |G|^{1/\rho} \leq |G|^\varepsilon \leq c(r, \varepsilon) \cdot |G|^{1/\varepsilon}.$$

For all graphs $G \in \mathcal{C}$ with $|G| < n_0(\rho)$, we trivially obtain the same bound

$$\text{wcol}_r(G(B)) \leq |G(B)| \leq |G| < n_0(\rho) = c(r, \varepsilon) \leq c(r, \varepsilon) \cdot |G|^{1/\varepsilon}. \quad \square$$

Lemma 17 states that graphs with small coloring numbers admit neighborhood covers with low degree. Combining Theorem 12 and Lemma 17 therefore directly gives us the following.

Lemma 35. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon)$ such that for every $G \in \mathcal{C}$, there exists an 8-contraction G' of G that admits an r -neighborhood cover with degree at most $c(r, \varepsilon) \cdot |G|^\varepsilon$ and spread at most $2r$.*

We now immediately obtain our main result with the help of Lemma 18. Flip-closed sparse neighborhood covers (with $\sigma(r)$ universally set to $34r$) follow by observing that each flipped class \mathcal{C}_ℓ is again structurally nowhere dense, and we may therefore apply the first part of Theorem 11 to it.

Theorem 11. *Let \mathcal{C} be a structurally nowhere dense class of graphs. For every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists $c(r, \varepsilon)$ such that for every $G \in \mathcal{C}$ there exists a weak r -neighborhood cover with degree at most $c(r, \varepsilon) \cdot |G|^\varepsilon$ and spread at most $34r$. In particular, \mathcal{C} admits flip-closed sparse weak neighborhood covers.*

References

- [AA14] Hans Adler and Isolde Adler. Interpreting nowhere dense graph classes as a classical notion of model theory. *European Journal of Combinatorics*, 36:322–330, 2014.
- [AT] **Anonymous Authors and Szymon Toruńczyk**. Nowhere dense model checking revisited. forthcoming.
- [BDG⁺22] Édouard Bonnet, Jan Dreier, Jakub Gajarský, Stephan Kreutzer, Nikolas Mählmann, Pierre Simon, and Szymon Toruńczyk. Model checking on interpretations of classes of bounded local cliquewidth. In Christel Baier and Dana Fisman, editors, *LICS '22: 37th Annual ACM/IEEE Symposium on Logic in Computer Science, Haifa, Israel, August 2 - 5, 2022*, pages 54:1–54:13. ACM, 2022.
- [BGdM⁺22] Édouard Bonnet, Ugo Giocanti, Patrice Ossona de Mendez, Pierre Simon, Stéphan Thomassé, and Szymon Toruńczyk. Twin-width IV: ordered graphs and matrices. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 924–937, 2022.
- [BKTW21] Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width I: tractable FO model checking. *ACM Journal of the ACM (JACM)*, 69(1):1–46, 2021.
- [BL22] Samuel Braufeld and Michael C Laskowski. Existential characterizations of monadic NIP. *arXiv preprint arXiv:2209.05120*, 2022.
- [BS85] John T Baldwin and Saharon Shelah. Second-order quantifiers and the complexity of theories. *Notre Dame Journal of Formal Logic*, 26(3):229–303, 1985.
- [CLS21] Michael B Cohen, Yin Tat Lee, and Zhao Song. Solving linear programs in the current matrix multiplication time. *Journal of the ACM (JACM)*, 68(1):1–39, 2021.
- [CMR00] Bruno Courcelle, Johann A Makowsky, and Udi Rotics. Linear time solvable optimization problems on graphs of bounded clique-width. *Theory of Computing Systems*, 33(2):125–150, 2000.
- [Cou90] Bruno Courcelle. The monadic second-order logic of graphs I. recognizable sets of finite graphs. *Information and computation*, 85(1):12–75, 1990.
- [DGK07] Anuj Dawar, Martin Grohe, and Stephan Kreutzer. Locally excluding a minor. In *22nd Annual IEEE Symposium on Logic in Computer Science (LICS 2007)*, pages 270–279. IEEE, 2007.

- [DGK⁺22a] Jan Dreier, Jakub Gajarský, Sandra Kiefer, Michał Pilipczuk, and Szymon Toruńczyk. Treelike decompositions for transductions of sparse graphs. In *Proceedings of the 37th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '22*, New York, NY, USA, 2022. Association for Computing Machinery.
- [DGK⁺22b] Jan Dreier, Jakub Gajarský, Sandra Kiefer, Michał Pilipczuk, and Szymon Toruńczyk. Treelike decompositions for transductions of sparse graphs, 2022.
- [Die12] Reinhard Diestel. *Graph Theory, 4th Edition*, volume 173 of *Graduate texts in mathematics*. Springer, 2012.
- [DKT10] Zdeněk Dvořák, Daniel Král, and Robin Thomas. Deciding first-order properties for sparse graphs. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 133–142. IEEE, 2010.
- [DMST22] Jan Dreier, Nikolas Mählmann, Sebastian Siebertz, and Szymon Toruńczyk. Indiscernibles and wideness in monadically stable and monadically NIP classes. *arXiv preprint arXiv:2206.13765*, 2022.
- [Dre21] Jan Dreier. Lacon- and shrub-decompositions: A new characterization of first-order transductions of bounded expansion classes. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021*, pages 1–13. IEEE, 2021.
- [Dvo07] Zdeněk Dvořák. Asymptotical structure of combinatorial objects, 2007.
- [Dvo22] Zdeněk Dvořák. Representation of short distances in structurally sparse graphs, 2022.
- [EF99] Heinz-Dieter Ebbinghaus and Jörg Flum. *Finite model theory*. Springer Science & Business Media, 1999.
- [EK17] Kord Eickmeyer and Ken-ichi Kawarabayashi. Fo model checking on map graphs. In *International Symposium on Fundamentals of Computation Theory*, pages 204–216. Springer, 2017.
- [Fef57] Solomon Feferman. Some recent work of Ehrenfeucht and Fraïssé. *Proc. of the Summer Institute of Symbolic Logic, Ithaca*, pages 201–209, 1957.
- [FG01a] Jörg Flum and Martin Grohe. Fixed-parameter tractability, definability, and model-checking. *SIAM Journal on Computing*, 31(1):113–145, 2001.
- [FG01b] Markus Frick and Martin Grohe. Deciding first-order properties of locally tree-decomposable structures. *Journal of the ACM (JACM)*, 48(6):1184–1206, 2001.
- [FG04] Markus Frick and Martin Grohe. The complexity of first-order and monadic second-order logic revisited. *Annals of pure and applied logic*, 130(1-3):3–31, 2004.
- [Gai82] Haim Gaifman. On local and non-local properties. In *Studies in Logic and the Foundations of Mathematics*, volume 107, pages 105–135. Elsevier, 1982.
- [GGK20] Jakub Gajarský, Maximilian Gorsky, and Stephan Kreutzer. Differential games, locality and model checking for FO logic of graphs. *arXiv preprint arXiv:2007.11345*, 2020.

- [GHK⁺13] Robert Ganian, Petr Hliněný, Daniel Král, Jan Obdržálek, Jarett Schwartz, and Jakub Teska. FO model checking of interval graphs. In *International Colloquium on Automata, Languages, and Programming*, pages 250–262. Springer, 2013.
- [GHL⁺14] Robert Ganian, Petr Hliněný, Alexander Langer, Jan Obdržálek, Peter Rossmanith, and Somnath Sikdar. Lower bounds on the complexity of mso1 model-checking. *Journal of Computer and System Sciences*, 80(1):180–194, 2014.
- [GHO⁺20] Jakub Gajarský, Petr Hliněný, Jan Obdržálek, Daniel Lokshtanov, and MS Ramanujan. A new perspective on FO model checking of dense graph classes. *ACM Transactions on Computational Logic (TOCL)*, 21(4):1–23, 2020.
- [GK11] Martin Grohe and Stephan Kreutzer. Methods for algorithmic meta theorems. *Model Theoretic Methods in Finite Combinatorics*, 558:181–206, 2011.
- [GKN⁺20] Jakub Gajarský, Stephan Kreutzer, Jaroslav Nešetřil, Patrice Ossona De Mendez, Michał Pilipczuk, Sebastian Siebertz, and Szymon Toruńczyk. First-order interpretations of bounded expansion classes. *ACM Transactions on Computational Logic (TOCL)*, 21(4):1–41, 2020.
- [GKR⁺18] Martin Grohe, Stephan Kreutzer, Roman Rabinovich, Sebastian Siebertz, and Konstantinos Stavropoulos. Coloring and covering nowhere dense graphs. *SIAM Journal on Discrete Mathematics*, 32(4):2467–2481, 2018.
- [GKS17] Martin Grohe, Stephan Kreutzer, and Sebastian Siebertz. Deciding first-order properties of nowhere dense graphs. *J. ACM*, 64(3):17:1–17:32, 2017.
- [GMM⁺23] Jakub Gajarský, Nikolas Mählmann, Rose McCarty, Pierre Ohlmann, Michał Pilipczuk, Wojciech Przybyszewski, Sebastian Siebertz, Marek Sokolowski, and Szymon Toruńczyk. Flipper games for monadically stable graph classes. *arXiv preprint arXiv:2301.13735*, 2023.
- [GPPT22] Jakub Gajarský, Michał Pilipczuk, Wojciech Przybyszewski, and Szymon Toruńczyk. Twin-Width and Types. In *49th International Colloquium on Automata, Languages, and Programming*, pages 123:1–123:21, 2022.
- [GPT22] Jakub Gajarský, Michał Pilipczuk, and Szymon Toruńczyk. Stable graphs of bounded twin-width. In *Proceedings of the 37th Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 1–12, 2022.
- [Gro08] Martin Grohe. Logic, graphs, and algorithms. *Logic and automata*, 2:357–422, 2008.
- [Hod97] Wilfrid Hodges. *A Shorter Model Theory*. Cambridge University Press, 1997.
- [HPR17] Petr Hliněný, Filip Pokrývka, and Bodhayan Roy. FO model checking of geometric graphs. In Daniel Lokshtanov and Naomi Nishimura, editors, *12th International Symposium on Parameterized and Exact Computation, IPEC 2017, September 6-8, 2017, Vienna, Austria*, volume 89 of *LIPICs*, pages 19:1–19:12. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.
- [IP01] Russell Impagliazzo and Ramamohan Paturi. On the complexity of k-SAT. *Journal of Computer and System Sciences*, 62(2):367–375, 2001.

- [IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Sciences*, 63(4):512–530, 2001.
- [Kre11] Stephan Kreutzer. Algorithmic meta-theorems. *Finite and algorithmic model theory*, 379:177–270, 2011.
- [KT10] Stephan Kreutzer and Siamak Tazari. Lower bounds for the complexity of monadic second-order logic. In *2010 25th Annual IEEE Symposium on Logic in Computer Science*, pages 189–198. IEEE, 2010.
- [KvRW⁺05] Fabian Kuhn, Pascal von Rickenbach, Roger Wattenhofer, Emo Welzl, and Aaron Zollinger. Interference in cellular networks: The minimum membership set cover problem. In Lusheng Wang, editor, *Computing and Combinatorics*, pages 188–198, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.
- [KY03] Hal A Kierstead and Daqing Yang. Orderings on graphs and game coloring number. *Order*, 20(3):255–264, 2003.
- [Lib04] Leonid Libkin. *Elements of finite model theory*, volume 41. Springer, 2004.
- [Mak04] Johann A Makowsky. Algorithmic uses of the Feferman–Vaught theorem. *Annals of Pure and Applied Logic*, 126(1-3):159–213, 2004.
- [NDM12] Jaroslav Nešetřil and Patrice Ossona De Mendez. *Sparsity: graphs, structures, and algorithms*, volume 28. Springer Science & Business Media, 2012.
- [NdM16] Jaroslav Nešetřil and P Ossona de Mendez. Structural sparsity. *Russian Mathematical Surveys*, 71(1):79, 2016.
- [Pel00] David Peleg. *Distributed computing: a locality-sensitive approach*. SIAM, 2000.
- [PPS20] Marcin Pilipczuk, Michał Pilipczuk, and Sebastian Siebertz. Lecture notes for the course “Sparsity” given at Faculty of Mathematics, Informatics, and Mechanics of the University of Warsaw, Winter semesters 2017/18 and 2019/20. Available <https://www.mimuw.edu.pl/~mp248287/sparsity2>.
- [PZ78] Klaus-Peter Podewski and Martin Ziegler. Stable graphs. *Fundamenta Mathematicae*, 100(2):101–107, 1978.
- [RT87] Prabhakar Raghavan and Clark D Tompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7(4):365–374, 1987.
- [See96] Detlef Seese. Linear time computable problems and first-order descriptions. *Mathematical Structures in Computer Science*, 6(6):505–526, 1996.
- [She90] Saharon Shelah. *Classification theory: and the number of non-isomorphic models*. Elsevier, 1990.
- [Tor22] Szymon Toruńczyk. Lectures on finite model theory. 2022.
- [TZ05] Mikkel Thorup and Uri Zwick. Approximate distance oracles. *Journal of the ACM (JACM)*, 52(1):1–24, 2005.

- [Vai89] Pravin M Vaidya. Speeding-up linear programming using fast matrix multiplication. In *30th annual symposium on foundations of computer science*, pages 332–337. IEEE Computer Society, 1989.
- [Vaz10] Vijay V. Vazirani. *Approximation Algorithms*. Springer Publishing Company, 2010.
- [war16] Algorithms, Logic and Structure Workshop in Warwick – Open Problem Session. https://warwick.ac.uk/fac/sci/math/people/staff/daniel_kral/alglogstr/openproblems.pdf, 2016. [Online; accessed 07-Feb-2023].