Chapter 2: Structural measures

Compilation date: November 12, 2017

1 Introduction

In the previous chapter we started by attempting to define sparsity by bounding the edge density, i.e., the ratio between the number of edges and the number of vertices, in the studied graph classes. This turned out to be equivalent (up to multiplicative factor 2 between the parameters) to bounding the *degeneracy*. Recall that a graph G is *d-degenerate* if every its subgraph has a vertex of maximum degree at most d, or equivalently if one can arrange the vertices of G into a linear order so that every vertex has at most d neighbors among vertices smaller in the order. The *degeneracy* of a graph is the minimum d for which this is possible. A vertex ordering with the minimum degeneracy provides sort of a decomposition for the graph, which can be algorithmically or combinatorially useful — e.g. for the purpose of employing some iteration or induction.

At the end of the day, we would like to study sparsity that is persistent under local contractions, as made explicit in the definitions of bounded expansion and nowhere denseness. Therefore, in our definitions we replaced the notion of a subgraph by the notion of a shallow minor, which enables us to inspect, in some sense, structures visible at any constant depth. It is natural to ask whether the definition of degeneracy via vertex orderings also admits a generalization to looking at any constant depth. The answer to this question is affirmative and comes in the form of *generalized coloring numbers*. These parameters, defined through the existence of vertex orderings with certain separation properties, are crucial tools for algorithmic and combinatorial treatment of classes of bounded expansion and, to some extent, as well of nowhere dense classes.

2 Definitions and basic properties

Let G be a graph. By a *vertex ordering* of G we mean any enumeration of $V(G)$ with numbers from 1 to $|V(G)|$, i.e., a bijective function $\sigma\colon V(G) \to \{1, \ldots, |V(G)|\}$. We often think of σ as the linear order $\leqslant_\sigma$ on the assigned precedences: for vertices $u, v \in V(G)$, we write $u \leqslant_\sigma v$ iff $\sigma(u) \leqslant \sigma(v)$.

We need to consider what is the right generalization of condition "every vertex has at most d neighbors among vertices smaller in the ordering" to looking at depth r instead of depth 0. There are three natural definitions, illustrated in Figure 1; each of them corresponds to a different generalized coloring number, and each of them is useful for some purposes. We first generalize, in two different ways, the concept of reaching a smaller vertex by a single edge to reaching a smaller vertex by a short path.

Definition 1. Let G be a graph, let σ be a vertex ordering of G, and let $r \in \mathbb{N}$. For vertices $u, v \in V(G)$ with $u \leqslant_\sigma v$, we say that:

- u is *strongly r-reachable* from v, if there is a path of length at most r from u to v whose every internal vertex w satisfies $v <_\sigma w$; and
- u is *weakly r-reachable* from v, if there is a path of length at most r from u to v whose every internal vertex w satisfies $u <_\sigma w$.

For a vertex v, the set of vertices strongly, respectively weakly, r-reachable from v in σ is denoted by $\mathrm{SReach}[G, \sigma, v]$, respectively by $\mathrm{WReach}[G, \sigma, v]$.

Note that every vertex is both weakly and strongly r-reachable from itself. Another notion of reaching is defined via the existence of many *disjoint* paths reaching smaller vertices.

Definition 2. Let G be a graph, let σ be a vertex ordering of G, and let $r \in \mathbb{N}$. The *r-admissibility* of a vertex v of G, denoted $\mathrm{adm}_r(G, \sigma, v)$, is equal to one plus the maximum size of a family of paths $\mathcal{P}$ with the following properties:

- every path from $\mathcal{P}$ has length at most r and leads from v to some vertex smaller than v in σ;
- paths from $\mathcal{P}$ are pairwise vertex-disjoint apart from sharing the endpoint v.

Observe that by trimming each path of $\mathcal{P}$ to the first encountered vertex smaller than v in σ, in the above definition we may assume without loss of generality that all the vertices traversed by paths from $\mathcal{P}$, apart from the endpoints other than v, are not smaller than v in σ. Note also that the r-admissibility is equal not to $|\mathcal{P}|$, where $\mathcal{P}$ is a path family as above, but to $1 + |\mathcal{P}|$. The rationale behind the $+1$ summand is to be consistent with the choice of definitions for weak and strong reachability; this will become clear later on.

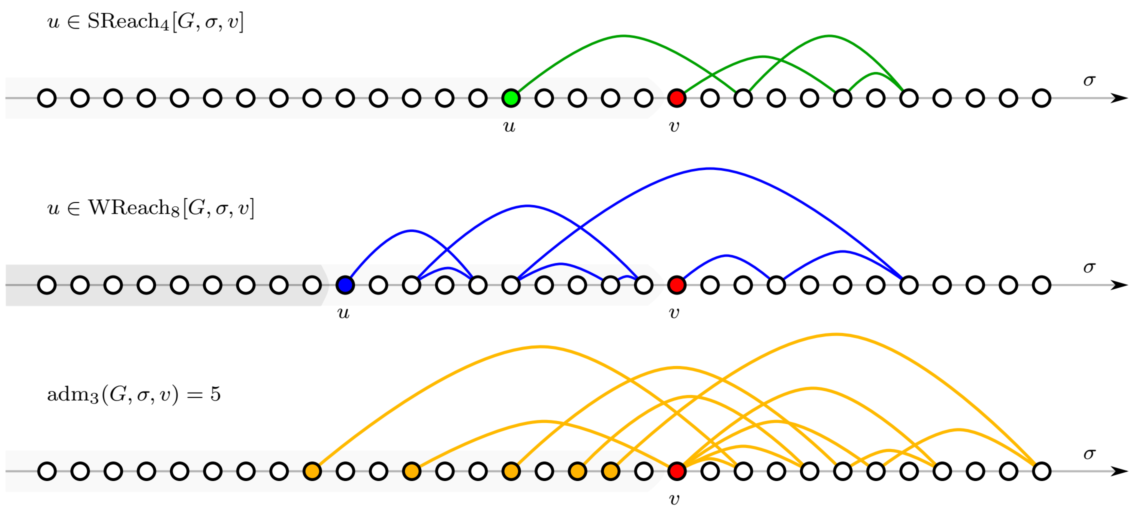

Figure 1: Different notions of reaching smaller vertices by short paths. In the first panel, u is strongly 4-reachable from v. In the second panel, u is weakly 8-reachable from v. In the third panel, the 3-admissibility of v is 5.

Note that $\mathrm{WReach}[G, v, \sigma]$ and $\mathrm{SReach}[G, v, \sigma]$ are sets, while $\mathrm{adm}_r(G, v, \sigma)$ is a number. We are now ready to define the generalized coloring numbers.

Definition 3. Let G be a graph and let $r \in \mathbb{N}$. For a vertex ordering σ of G, we define the *weak r-coloring number*, the *strong r-coloring number*, and the *r-admissibility* of σ as follows:

$$\begin{aligned} \mathrm{wcol}_r(G, \sigma) &\coloneqq \max_{v \in V(G)} |\mathrm{WReach}[G, \sigma, v]|, \\ \mathrm{col}_r(G, \sigma) &\coloneqq \max_{v \in V(G)} |\mathrm{SReach}[G, \sigma, v]|, \\ \mathrm{adm}_r(G, \sigma) &\coloneqq \max_{v \in V(G)} \mathrm{adm}(G, \sigma, v). \end{aligned}$$

The *weak r-coloring number*, the *strong r-coloring number*, and the *r-admissibility* of G are defined as the minimum among vertex orderings σ of G of the respective parameter for σ. That is, if by $\Pi(G)$ we denote the set of vertex orderings of G, then

$$\begin{aligned} \mathrm{wcol}_r(G) &:= \min_{\sigma\in\Pi(G)} \mathrm{wcol}_r(G,\sigma), \\ \mathrm{col}_r(G) &:= \min_{\sigma\in\Pi(G)} \mathrm{col}_r(G,\sigma), \\ \mathrm{adm}_r(G) &:= \min_{\sigma\in\Pi(G)} \mathrm{adm}_r(G,\sigma). \end{aligned}$$

Note that for $r = 1$, all the above three notions are equal to the degeneracy plus one; for $r > 1$ the notions are already different. The following inequalities follow directly from the definitions.

Proposition 1. *For every $r \in \mathbb{N}$, graph G, and its vertex ordering σ, the following holds:*

$$\mathrm{adm}_r(G,\sigma) \leqslant \mathrm{col}_r(G,\sigma) \leqslant \mathrm{wcol}_r(G,\sigma).$$

Proof. For the second inequality, note that $\mathrm{SReach}[G,\sigma,v] \subseteq \mathrm{WReach}[G,\sigma,v]$ for all $v \in V(G)$. For the first inequality, note that if for a vertex v we have a path family $\mathcal{P}$ witnessing the value of $\mathrm{adm}_r(G,\sigma,v)$, and without loss of generality the endpoints of paths from $\mathcal{P}$ other than v are the only vertices traversed by these paths that are smaller than v in σ, then each of these endpoints belongs to $\mathrm{SReach}[G,\sigma,v] - \{v\}$. □

It appears that the generalized coloring numbers are actually functionally equivalent: we not only have the bounds as in Proposition 1, but actually any of them can be bounded both from below and from above by a function of any other one. This enables convenient switching between coloring numbers according to which is more suitable for a particular need. We prove this fact in the following two lemmas.

Lemma 2. *For every $r \in \mathbb{N}$, graph G, and its vertex ordering σ, the following holds:*

$$\mathrm{col}_r(G,\sigma) \leqslant 1 + (\mathrm{adm}_r(G,\sigma) - 1)^r.$$

Proof. Let $k := \mathrm{adm}_r(G,\sigma) - 1$; we need to prove that $\mathrm{col}_r(G,\sigma) \leqslant 1 + k^r$. Take any vertex v and let $A := \{u \colon u <_\sigma v\}$ be the set of vertices smaller than v in σ. Run a BFS from v in the graph $G - A$, and let T be the obtained shortest path spanning tree of the connected component of v in $G - A$. For each vertex $u \in A$ that is strongly r-reachable from u in σ, there is a path P_u of length at most r witnessing this fact. Path P_u starts in v, and traverses only vertices of $V(G) - A$ before finally jumping to $u \in A$ and finishing there. If $u' \in V(G) - A$ is the vertex preceding u on P_u, then we can replace the subpath of P_u from v to u' by a shortest path between v and u' contained in T; the length of P_u does not increase in this manner. Thus, from now on we may assume that all paths P_u are contained in the tree T, apart from their last edges.

Let T' be the union of paths P_u for $u \in \mathrm{SReach}[G,\sigma,v] - \{v\}$; see Figure 2 for an illustration. From the above assumption it follows that T' is a tree. Moreover, if we root T' at v, then T' has depth at most r and the leaves of T' are exactly the vertices of $\mathrm{SReach}[G,\sigma,v] - \{v\}$.

We now claim that every internal vertex w of T' has at most k children in T'. Indeed, otherwise we could find more than k vertex-disjoint paths of length at most r from w to A, one per each subtree rooted at a child of w. Since $w \notin A$ implies $v \leqslant_\sigma w$, this would imply $\mathrm{adm}_r(G,\sigma,w) > k+1$, contradicting the assumption that $\mathrm{adm}_r(G,\sigma) = k + 1$.

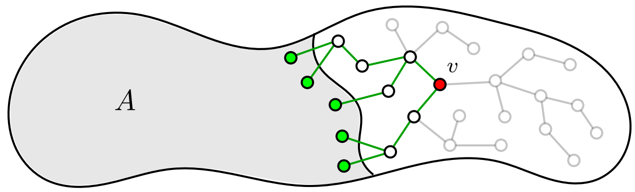

Figure 2: Construction of the tree T' (in green). The remainder of the tree T is depicted in grey.

Summarizing, T' is a tree of depth at most r, whose every internal vertex has at most k children. Hence T' has at most k^r leaves. Since the leaves of T' are exactly the vertices of $\mathrm{SReach}[G,\sigma,v]-\{v\}$, it follows that $|\mathrm{SReach}[G,\sigma,v]| \leqslant 1+k^r$; as v was chosen arbitrarily, this concludes the proof. $\square$

Lemma 3. *For every $r \in \mathbb{N}$, graph G, and its vertex ordering σ, the following holds:*

$$\mathrm{wcol}_r(G,\sigma) \leqslant 1 + r(\mathrm{col}_r(G,\sigma)-1)^r.$$

Proof. Let $k \coloneqq \mathrm{col}_r(G,\sigma)-1$; we need to prove that $\mathrm{wcol}_r(G,\sigma) \leqslant 1+rk^r$. For each vertex u consider the set $B_u \coloneqq \mathrm{SReach}_r[G,\sigma,u]-\{u\}$, which obviously has size at most k, and fix an arbitrary enumeration $\{b_u^1, b_u^2, \ldots, b_u^\ell\}$ of B_u, where $\ell = |B_u| \leqslant k$.

Let us fix an arbitrary vertex v of G; we are going to examine its weak r-reachability set. Take any u that belongs to $\mathrm{WReach}_r[G,\sigma,v]-\{v\}$ and let P be any path witnessing this fact. That is, P has length at least 1 and at most r, leads from v to u, and all internal vertices of P are not smaller than u in σ. Call a vertex w on P a *milestone* if all vertices traversed by P between v and w are not smaller than w in σ (see Figure 3 for an example). Note that v itself is the first milestone, whereas u is the last milestone by the definition of weak reachability. Let $w_1, w_2, \ldots, w_p$ be the consecutive milestones on P, where $w_1 = v$ and $w_p = u$. It is straightforward to see from the definition that

$$u = w_p <_\sigma w_{p-1} <_\sigma \ldots <_\sigma w_2 <_\sigma w_1 = v.$$

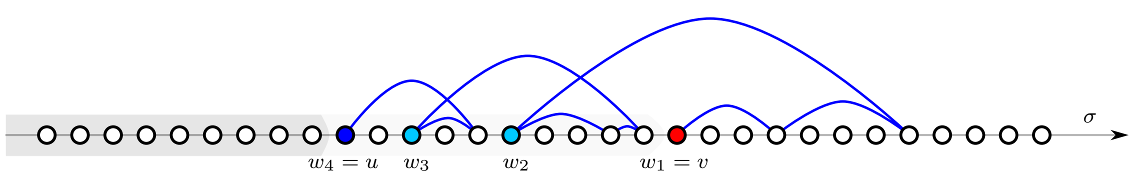

Figure 3: The example path from the second panel of Figure 1 has 4 milestones: $w_1 = u$, $w_2 = v$, and two intermediate ones.

We now claim the following: for each $i \in \{1,\ldots,p-1\}$, the infix of P between w_i and w_{i+1} witnesses that $w_{i+1} \in \mathrm{SReach}_r[G,\sigma,w_i]-\{w_i\}$. Indeed, this infix has length at most r, and all vertices traversed by P between w_i and w_{i+1} (exclusive) have to be larger than w_i in σ, for otherwise we would see another milestone between w_i and w_{i+1}.

Since $w_{i+1} \in \mathrm{SReach}_r[G,\sigma,w_i]-\{w_i\} = B_{w_i}$, we have that $w_{i+1} = b_{w_i}^{j_i}$ for some j_i satisfying $1 \leqslant j_i \leqslant |B_{w_i}| \leqslant k$. Now define the *signature* of u as follows:

$$\phi(u) \coloneqq (j_1, j_2, \ldots, j_{p-1}).$$

As the length of P is at most r, we have $p \leqslant r+1$. Therefore, the signature of any vertex $u \in \mathrm{WReach}_r[G,\sigma,v]-\{v\}$ is a nonempty sequence of length at most r with entries from $\{1,\ldots,k\}$.

Now comes the crucial observation: different vertices $u,u' \in \mathrm{WReach}_r[G,\sigma,v]-\{v\}$ always receive different signatures; that is, for $u \neq u'$ we have $\phi(u) \neq \phi(u')$. To see this, observe that starting with the first milestone v, one can, using a straightforward induction, recover consecutive milestones from the signature alone. The last milestone is the vertex whose signature we consider.

We infer that the size of $\mathrm{WReach}_r[G,\sigma,v]-\{v\}$ is bounded by the number of nonempty sequences of length at most r with entries from $\{1,\ldots,k\}$, which in turn is equal to $k^1+k^2+\ldots+k^r \leqslant rk^r$. Thus $|\mathrm{WReach}_r[G,\sigma,v]| \leqslant 1+rk^r$ for every vertex v, implying that $\mathrm{wcol}_r(G,\sigma) \leqslant 1+rk^r$. □

Putting Proposition 1 and Lemmas 2 and 3 together yields the following.

Corollary 4. *For every $r \in \mathbb{N}$, graph G, and its vertex ordering σ, the following holds:*

$$\mathrm{adm}_r(G,\sigma) \leqslant \mathrm{col}_r(G,\sigma) \leqslant \mathrm{wcol}_r(G,\sigma) \leqslant 1+r(\mathrm{adm}_r(G,\sigma)-1)^{r^2}.$$

In particular, for every $r \in \mathbb{N}$ and graph G we have:

$$\mathrm{adm}_r(G) \leqslant \mathrm{col}_r(G) \leqslant \mathrm{wcol}_r(G) \leqslant 1+r(\mathrm{adm}_r(G)-1)^{r^2}.$$

We will later use the fact that admissibility has good algorithmic properties: it is relatively easy to compute it. More precisely, in future lectures we will use the following statement.

Theorem 5. *Let $\mathcal{C}$ be a fixed class of bounded expansion and let $r \in \mathbb{N}$ be fixed. Then there exists an algorithm that, given an n-vertex graph $G \in \mathcal{C}$, computes a vertex ordering σ of G with $\mathrm{adm}_r(G,\sigma) = \mathrm{adm}_r(G)$ in time $\mathcal{O}(n)$.*

Since $\mathcal{C}$ and r are fixed in the above statement, the constants hidden in the $\mathcal{O}(\cdot)$ notation may, and do, depend on r and the parameters (grads) of $\mathcal{C}$; this dependence is exponential. The proof of Theorem 5 is quite difficult and relies on auxiliary data structures, and therefore we will not cover it. However, during the tutorials we will see a simple proof of the following result.

Theorem 6. *There is an algorithm that given a graph G with n vertices and m edges, and parameter $r \in \mathbb{N}$, computes in time $\mathcal{O}(n^3m)$ a vertex ordering σ of G with $\mathrm{adm}_r(G,\sigma) \leqslant r \cdot \mathrm{adm}_r(G)$.*

The algorithm of Theorem 6 can be applied on any graph, and the constant hidden in the $\mathcal{O}(\cdot)$ notation does not depend on r (i.e., r is not fixed). In general most of the algorithms presented in this chapter can be made to work in linear time (with hidden multiplicative constants heavily depending on r and the class of bounded expansion from which a graph is drawn), but in order not to obscure the presentation with implementation details of secondary importance, we will omit these aspects.

Observe that the abovementioned algorithms for computing (approximate) r-admissibility may also serve as approximation algorithms for the weak and the strong r-coloring numbers, due to the bounds of Corollary 4. This is particularly important for the algorithmic aspects of the theory of sparse graphs, as many algorithmic results rely on first computing a vertex ordering with bounded weak r-coloring number, and then using it for further computations. Developing efficient approximation algorithms for the weak r-coloring number is a major open problem, both from the theoretical and from the practical point of view.

3 Relation with density of shallow minors

The crucial point of the generalized coloring numbers is that they are not only functionally equivalent to each other, but they are also functionally equivalent to the density of shallow minors. Therefore, classes of bounded expansion may be equivalently defined as those, where for every $r \in \mathbb{N}$ the weak r-coloring numbers of graphs from the class are bounded by a constant depending only on r. We prove this in the following two lemmas. The first one is easy: a dense depth-r minor witnesses that there is no vertex ordering with a small weak $(4r+1)$-coloring number. These second one will be harder: we will prove that if an algorithm for computing the r-admissibility of a graph fails to produce an ordering with small admissibility, this is because it encounters an obstacle in the form of a dense depth-$(r-1)$ topological minor.

Lemma 7. *For every $r \in \mathbb{N}$ and graph G the following holds:*

$$\nabla_r(G) \leqslant \mathrm{wcol}_{4r+1}(G).$$

Proof. Let $d \coloneqq \mathrm{wcol}_{4r+1}(G)$. It suffices to show that every depth-r minor H of G contains a vertex of degree at most d. Indeed, since the class $\{G\}\nabla r$ is closed under taking subgraphs by definition, this would imply that every depth-r minor of G is d-degenerate, which means that the ratio between the number of its edges and the number of its vertices is at most d.

Let H be a depth-r minor of G and let ϕ be a depth-r minor model witnessing this fact. Further, let σ be a vertex ordering of G witnessing that $\mathrm{wcol}_{4r+1}(G) \leqslant d$; that is, $|\mathrm{WReach}_{4r+1}[G, \sigma, v]| \leqslant d$ for each $v \in V(G)$. For every vertex $u \in V(H)$, let $\gamma(u)$ be the smallest, in σ, vertex of the branch set $\phi(u)$; since the branch sets are pairwise disjoint, vertices $\phi(u)$ are pairwise different for $u \in V(H)$. Let then $u_{\max} \in V(H)$ be the vertex of H for which $\gamma(u_{\max})$ is the largest in σ. We claim that $u_{\max}$ has at most d neighbors in H, which will conclude the proof.

Take any neighbor w of $u_{\max}$ in H. Since the branch sets $\phi(u_{\max})$ and $\phi(w)$ have radii at most r and there is an edge between them, there is a path P of length at most $4r+1$ that leads from $\gamma(u_{\max})$ to $\gamma(w)$, and which traverses only vertices of $\phi(u_{\max}) \cup \phi(w)$. Observe further that $\gamma(w)$ is smaller in σ than all the other vertices of $\phi(u_{\max}) \cup \phi(w)$: it is smaller than all the other vertices of $\phi(w)$ by definition, and by the choice of $u_{\max}$ it is also smaller than $\gamma(u_{\max})$, which in turn is smaller than all the other vertices of $\phi(u_{\max})$. This means that P witnesses that $\gamma(w) \in \mathrm{WReach}_{4r+1}[G, \sigma, \gamma(u_{\max})]$. Since this weak reachability set has size at most d, we conclude that indeed $u_{\max}$ has at most d neighbors in H. □

Lemma 8. *For every $r \in \mathbb{N}$ and graph G, the following holds*

$$\mathrm{adm}_r(G) \leqslant 1 + 6r \left(\lceil \widetilde{\nabla}_{r-1}(G) \rceil \right)^3.$$

Proof. For a set $S \subseteq V(G)$ and $v \in S$, let $b_r(S, v)$ be the maximum size of a family $\mathcal{P}$ of paths in G with the following properties:

- each path $P \in \mathcal{P}$ has length at most r, leads from v to some other vertex of S, and all its internal vertices do not belong to S; and
- for all distinct $P, P' \in \mathcal{P}$, we have $V(P) \cap V(P') = \{v\}$.

We order the vertices of G as $v_1, v_2, \ldots, v_n$ as follows. Assume $v_{i+1}, \ldots, v_n$ have already been ordered. Define $S_i \coloneqq \{v_1, \ldots, v_i\}$. Choose any $v \in S_i$ such $b_r(S_i, v)$ is minimum possible, and define $v_i \coloneqq v$. (In particular, v_n is any vertex of minimum degree in G.) Clearly, the r-admissibility of the resulting order is $1 + \max_{1 \leqslant i \leqslant n} b_r(S_i, v_i)$.

Let $d \coloneqq \lceil \widetilde{\nabla}_{r-1}(G) \rceil$ and assume towards a contradiction that in the above construction we encounter in some iteration i a set $S \coloneqq S_i$ such that $b_r(S, v) > \ell \coloneqq 6rd^3$ for all $v \in S$. For each $v \in S$, fix a family of paths $\mathcal{P}_v$ witnessing that $b_r(S, v) > \ell$; in particular, $|\mathcal{P}_v| > \ell$. Let $s \coloneqq |S|$.

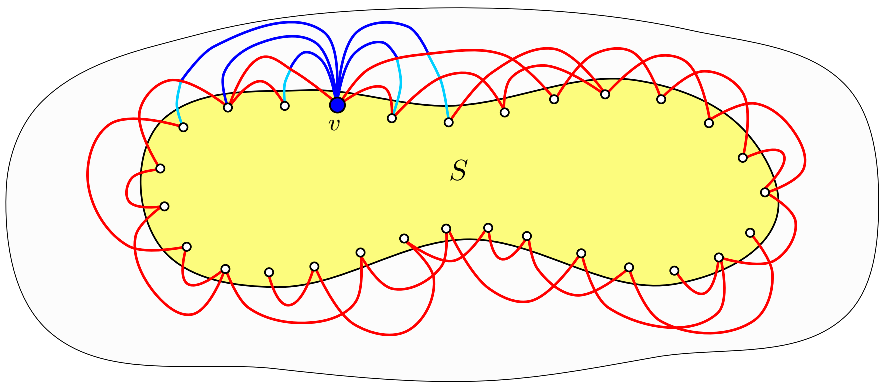

Figure 4: Situation in the proof of Lemma 8. Paths from $\mathcal{Q}$ are depicted in red; even though this might not be visible in the figure, they are pairwise internally vertex-disjoint. For one particular vertex v, families $\mathcal{P}_v$ and $\mathcal{P}'_v$ are depicted. The latter one is in dark blue, while the suffixes of paths from $\mathcal{P}_v$ that were dropped in the construction of paths from $\mathcal{P}'_v$ are in light blue.

Our goal is to construct from the set S and path families $\{\mathcal{P}_v\}_{v \in S}$ a depth-$(r-1)$ topological minor of G which is too dense, i.e., has edge density larger than d. As a first step, we choose a maximal family $\mathcal{Q}$ of paths satisfying the following conditions:

- each path from $\mathcal{Q}$ has length at most $2r-1$, connects two distinct vertices of S, and all its internal vertices do not belong to S;
- each pair of distinct vertices of S is connected by at most one path from $\mathcal{Q}$; and
- paths from $\mathcal{Q}$ are pairwise internally vertex-disjoint.

Note that these are not paths from the families $\mathcal{P}_v$, but arbitrary paths in G. Let H be the graph with vertex set S and edges between all pairs of vertices $u, v \in S$ that are connected by a path in $\mathcal{Q}$. Then $H \preccurlyeq^{\mathrm{top}}_{r-1} G$, hence $|\mathcal{Q}| = |E(H)| \leqslant d \cdot s$. Let K be the set of all internal vertices of the paths in $\mathcal{Q}$. As every path from $\mathcal{Q}$ has at most $2r-2$ internal vertices, we have $|K| \leqslant s \cdot d \cdot (2r-2)$.

As $H \preccurlyeq^{\mathrm{top}}_{r-1} G$, we have that H is $2d$-degenerate. Hence H admits a proper coloring with $(2d+1)$ colors, which implies that H contains an independent set I of size at least $\frac{s}{2d+1}$.

For every $v \in S$, we define $\mathcal{P}'_v$ to be the family of prefixes paths in $\mathcal{P}_v$ from v to a vertex in $(S \cup K) - \{v\}$ with all internal vertices in $V(G) - (S \cup K)$. More precisely, for every path $P \in \mathcal{P}_v$, we let P' be the prefix of P from v to the first vertex belonging to $(S \cup K) - \{v\}$, and we let $\mathcal{P}'_v$ comprise paths P' for all $P \in \mathcal{P}_v$. Note that P' is correctly defined because the endpoint of P different from v belongs to $S - \{v\}$. Obviously, we have $|\mathcal{P}'_v| = |\mathcal{P}_v|$ for all $v \in S$.

Here comes the crucial observation: for distinct $u, v \in I$, the paths in $\mathcal{P}'_u$ and $\mathcal{P}'_v$ are pairwise internally vertex-disjoint. Indeed, suppose that some paths $P_1 \in \mathcal{P}'_u$ and $P_2 \in \mathcal{P}'_v$ intersected at some vertex $w \in V(G) - (K \cup S)$. Then the union of paths P_1 and P_2 would contain a path of length at most $2r - 2$ connecting u and v that is internally disjoint from all paths in $\mathcal{Q}$. Since u and v are not adjacent in H (recall that $u, v \in I$ and I is an independent set in H), this path could be added to $\mathcal{Q}$. This would contradict the maximality of $\mathcal{Q}$.

We now construct a depth-$(r-1)$ topological minor J of G on vertex set $S \cup K$ as follows: contract all paths in $\bigcup_{v \in I} \mathcal{P}'_v$ to single edges. By the observation of the previous paragraph, all these paths are pairwise internally vertex-disjoint, so this contraction is well-defined and yields a depth-$(r-1)$ topological minor of G.

It remains to estimate the edge density of J. On one hand, we have

$$|V(J)| \leqslant |S| + |K| \leqslant s + s \cdot d \cdot (2r-2) \leqslant s \cdot d \cdot (2r-1).$$

On the other hand, every vertex $v \in I$ brings at least $|\mathcal{P}'_v| > \ell$ edges to J, hence

$$|E(J)| > |I| \cdot \ell \geq \frac{s}{2d+1} \cdot \ell.$$

The lemma statement is trivial when G is edgeless, so we may assume otherwise; in particular $d \geqslant 1$. Then

$$\frac{|E(J)|}{|V(J)|} > \frac{s \cdot \ell}{(2d+1) \cdot s \cdot d \cdot (2r-1)} > \frac{\ell}{6rd^2} = d,$$

which is a contradiction with J being a depth-$(r-1)$ topological minor of G. □

Lemmas 7 and 8 together with the results from Chapter 1 (more precisely, the bounds between densities of shallow minors and shallow topological minors) yield the following.

Theorem 9. *Let $\mathcal{C}$ be a class of graphs. Then the following conditions are equivalent.*

(a) $\mathcal{C}$ has bounded expansion.

(b) There is a function $f \colon \mathbb{N} \to \mathbb{N}$ such that $\mathrm{wcol}_r(G) \leqslant f(r)$ for all $G \in \mathcal{C}$ and $r \in \mathbb{N}$.

(c) There is a function $f \colon \mathbb{N} \to \mathbb{N}$ such that $\mathrm{col}_r(G) \leqslant f(r)$ for all $G \in \mathcal{C}$ and $r \in \mathbb{N}$.

(d) There is a function $f \colon \mathbb{N} \to \mathbb{N}$ such that $\mathrm{adm}_r(G) \leqslant f(r)$ for all $G \in \mathcal{C}$ and $r \in \mathbb{N}$.

Further, in the first chapter we proved that if a class $\mathcal{C}$ is nowhere dense, then there is a function $f \colon \mathbb{N} \times \mathbb{R} \to \mathbb{N}$ such that for all $r \in \mathbb{N}$, $\varepsilon > 0$, and $G \in \mathcal{C} \nabla r$, we have that $\frac{|E(G)|}{|V(G)|} \leqslant f(r, \varepsilon) \cdot |V(G)|^{\varepsilon}$. Observe that the relations between parameters: grads, topological grads, weak coloring numbers, strong coloring numbers, admissibility, are governed by polynomial upper bounds; that is, each parameter above is bounded by a polynomial of any other parameter, possibly with increased radius r. This yields the following.

Theorem 10. *Let $\mathcal{C}$ be nowhere dense class of graphs. Then there is a function $f \colon \mathbb{N} \times \mathbb{R} \to \mathbb{N}$ such that $\mathrm{wcol}_r(G) \leqslant f(r, \varepsilon) \cdot |V(G)|^{\varepsilon}$ for all $r \in \mathbb{N}$, $\varepsilon > 0$, and $G \in \mathcal{C}$.*

4 Duality between independent sets and dominating sets

In this section we provide one of the foremost examples of applications of generalized coloring numbers: approximation algorithms for the independence number and the domination number of a graph. We start with a few definitions.

For a radius $r \in \mathbb{N}$ and a graph G, a subset of vertices $I \subseteq V(G)$ is *r-independent* if vertices of I are pairwise at distance more than r; formally, for all distinct $u, v \in I$, we have $\mathrm{dist}_G(u, v) > r$. Note that a vertex subset I is $2r$-independent in G if and only if the r-neighborhoods (balls of radius r) of vertices of I are pairwise disjoint. The size of a largest r-independent set in a graph G is called the *r-indepdendence number* of G and will be denoted by $\mathrm{ind}_r(G)$.

A subset of vertices $D \subseteq V(G)$ is called an *r-dominating set* in G if every vertex of G is at distance at most r from some vertex of D. Equivalently, the r-neighborhoods of vertices of D in total cover the whole vertex set of G. The size of a smallest r-dominating set in a graph G is called the *r-domination number* of G and will be denoted by $\mathrm{dom}_r(G)$.

Thus, finding large $2r$-independent sets is sort of a packing problem, where we try to pack as many disjoint balls of radius r in a given graph, while finding small r-dominating sets is sort of a covering problem, where we want to cover the graph with as few balls of radius r as possible. It is easy to see that in any graph, the optimum for the first problem is a lower bound on the optimum for the second.

Lemma 11. *For every $r \in \mathbb{N}$ and graph G, the following holds:*

$$\mathrm{ind}_{2r}(G) \leqslant \mathrm{dom}_r(G).$$

Proof. Let I be a $2r$-independent set of maximum size in G. Then r-neighborhoods of vertices of I are pairwise disjoint, hence every r-dominating set in G has to contain at least one vertex from each of these r-neighborhoods in order to r-dominate I. Consequently, every r-dominating set of G has size at least $|I|$. □

In general, there are graphs with 2-independence number equal to 1 and arbitrarily large 1-domination number, so we cannot hope for any reasonable inequality in the other direction. However, if we restrict attention to classes of bounded expansion, it turns out that between the $2r$-independence number and r-domination number there is a constant multiplicative gap.

Theorem 12. *For every $r \in \mathbb{N}$ and graph G, the following holds:*

$$\mathrm{dom}_r(G) \leqslant \mathrm{wcol}_{2r+1}(G)^2 \cdot \mathrm{ind}_{2r+1}(G).$$

Moreover, there is an algorithm with running time $\mathcal{O}(n^3 m)$ that given G, r, and a vertex ordering σ of G with $\mathrm{wcol}_{2r+1}(G, \sigma) = c$, computes an r-dominating set D of G and a $(2r+1)$-independent set I of G satisfying $|D| \leqslant c^2 \cdot |I|$.

Note that in Theorem 12, the right hand side of the inequality contains the $(2r+1)$-independence number, which may be even smaller than the $2r$-independence number. Before we proceed to the proof of Theorem 12, let us infer the following algorithmic corollary. Here, we consider the algorithmic problems r-Dominating Set and r-Independent Set, where given a graph G and parameter $r \in \mathbb{N}$, our goal is to compute the smallest r-dominating set, respectively the largest r-independent set, in G.

Corollary 13. *For every class of bounded expansion $\mathcal{C}$ and every $r \in \mathbb{N}$, the r-*Dominating Set *problem and the r-*Independent Set *problem admit constant-factor approximation algorithms running in time $\mathcal{O}(n^3 m)$. The approximation factor depends on $\mathcal{C}$ and r, while the constant hidden in the $\mathcal{O}(\cdot)$ notation does not.*

Proof. By Theorem 6, given a graph $G \in \mathcal{C}$ and $r \in \mathbb{N}$ we may compute in time $\mathcal{O}(n^3 m)$ a vertex ordering σ with $\mathrm{adm}_{2r+1}(G, \sigma) \leqslant (2r+1) \cdot \mathrm{adm}_{2r+1}(\mathcal{C})$. By Corollary 4, we have

$$\mathrm{wcol}_{2r+1}(G, \sigma) \leqslant 1 + (2r+1)((2r+1) \cdot \mathrm{adm}_{2r+1}(\mathcal{C}) - 1)^{(2r+1)^2} =: c.$$

Now, apply the algorithm of Theorem 12 on G, r, and σ, yielding an r-dominating set D and a $(2r+1)$-independent set I with $|D| \leqslant c^2 \cdot |I|$. By Lemma 11, we have

$$|D| \leqslant c^2 \cdot |I| \leqslant c^2 \cdot \mathrm{ind}_{2r+1}(G) \leqslant c^2 \cdot \mathrm{dom}_r(G),$$

so the size of D is at most c^2 times the optimum. This yields the approximation algorithm for the r-Dominating Set problem. For the independence, observe that

$$|I| \geqslant |D|/c^2 \geqslant \mathrm{dom}_r(G)/c^2 \geqslant \mathrm{ind}_{2r}(G)/c^2 \geqslant \mathrm{ind}_{2r+1}(G)/c^2,$$

so the size of I is at least $1/c^2$ times the optimum, both for $2r$-independent sets and for $(2r+1)$-independent sets. This yields the approximation algorithm for $2r$-Independent Set and $(2r+1)$-Independent Set. □

We now proceed to the proof of Theorem 12.

Proof of Theorem 12. We give a proof of the algorithmic result, as it trivially implies the stated inequality by taking σ to be a vertex ordering of G with the optimum $(2r+1)$-weak coloring number. It is easy, using a breadth-first search, to verify in time $\mathcal{O}(m)$ for two vertices $u <_\sigma v$ whether $u \in \mathrm{WReach}_{2r+1}[G, \sigma, v]$. By applying this for every pair of vertices, we may compute in time $\mathcal{O}(n^2 m)$ the set $\mathrm{WReach}_{2r+1}[G, \sigma, v]$ for every $v \in V(G)$.

We now apply the following greedy procedure, which will maintain three sets of vertices: D, the constructed r-dominating set; A, which will eventually be turned into a $(2r+1)$-independent set; and R, the set of vertices that remain to be dominated. We maintain the invariant that R comprises vertices of G that are not r-dominated by D.

1. Start with $D := \emptyset$, $A := \emptyset$, and $R := V(G)$.
2. As long as R is non-empty, perform the following:
   (a) Let v be the vertex of R that is the smallest in σ.
   (b) Add all vertices of $\mathrm{WReach}_{2r+1}[G, \sigma, v]$ to D.
   (c) Remove from R every vertex that became r-dominated by a vertex added to D.

It is straightforward to implement the block under the loop in time $\mathcal{O}(nm)$, by running a breadth-first search from every vertex added to D, so the whole algorithm runs in time $\mathcal{O}(n^2 m)$. Also, the following assertions follow immediately from the algorithm:

- At the end D is an r-dominating set of G, because R becomes empty.

- At the end we have $|D| \leqslant c \cdot |A|$, as with every vertex added to A we add at most c vertices to D.

It remains to find a large $(2r+1)$-independent set within A. We do this using the following two claims.

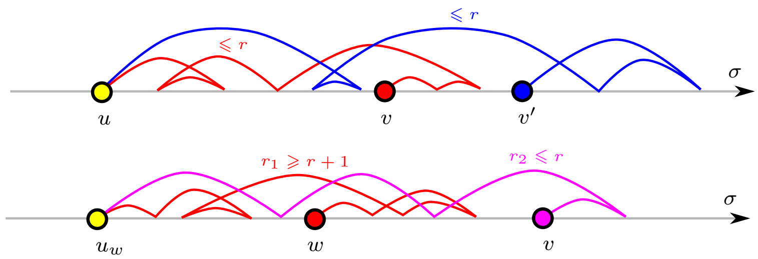

Figure 5: Illustration for the two claims in the proof of Theorem 12.

Claim 1. *For each $u \in V(G)$ there is at most one vertex $v \in A$ such that $u \in \mathrm{WReach}_r[G,\sigma,v]$.*

Proof. Suppose otherwise: there are two different vertices $v, v' \in A$ with $u \in \mathrm{WReach}_r[G,\sigma,v] \cap \mathrm{WReach}_r[G,\sigma,v']$. Without loss of generality suppose $v <_\sigma v'$; this implies that v was added to A before v'. Since $u \in \mathrm{WReach}_r[G,\sigma,v]$, we have that u was added to D when v was added to A (or u was added to D even before). In particular, at this moment, every vertex at distance at most r from u, in particular v', was removed from R. However, this contradicts the assumption that $v' \in A$, since we select to A only vertices that, at the moment of their selection, belong to R. ⌟

Claim 2. *For each $v \in A$ there are at most c vertices $w \in A$ with $w \leqslant_\sigma v$ and $\mathrm{dist}_G(v,w) \leqslant 2r+1$.*

Proof. Take any such vertex $w \in A$ and let P be any path connecting w and v that has length at most $2r+1$. Let u_w be the vertex of P that is the smallest in σ, and let r_1, r_2 be the lengths of the subpaths of P from w to u_w and from u_w to v, respectively; in particular $r_1 + r_2 \leqslant 2r+1$. These subpaths certify that $u_w \in \mathrm{WReach}_{2r+1}[G,\sigma,w] \cap \mathrm{WReach}_{2r+1}[G,\sigma,v]$, so in particular u_w was added to D when w was added to A (or even before). Observe that we cannot have $r_2 \leqslant r$, because then after adding u_w to D the vertex v would be removed from R, a contradiction with $v \in A$. Hence $r_2 \geqslant r+1$, so $r_1 \leqslant r$, and thus $u_w \in \mathrm{WReach}_r[G,\sigma,w]$. By the first claim we infer that w is the unique vertex of A for which $u_w \in \mathrm{WReach}_r[G,\sigma,w]$. We conclude that vertices u_w are pairwise different for all $w \in A$ with $w \leqslant_\sigma v$ and $\mathrm{dist}_G(v,w) \leqslant 2r+1$. Since every such vertex u_w belongs to $\mathrm{WReach}_{2r+1}[G,\sigma,v]$ and this set has size at most c, the claim follows. ⌟

Construct now an auxiliary graph H on the vertex set A, where distinct vertices $v, w \in A$ are adjacent if and only if $\mathrm{dist}_G(v,w) \leqslant 2r+1$. It is straightforward to construct H in time $\mathcal{O}(nm)$ by running a breadth-first search from every vertex of A. The second claim states that the restriction of σ to A is a vertex ordering of H with degeneracy $c-1$, which implies that H admits a proper coloring with c colors; such a coloring can be obtained in time $\mathcal{O}(n^2)$ by employing a greedy algorithm. At least one of the color classes in this coloring has size at least $|A|/c$; let us denote it by I. By construction we have that I is a $(2r+1)$-independent set in G and it has size

$$|I| \geqslant |A|/c \geqslant |D|/c^2,$$

which concludes the proof. □

5 Low tree-depth decompositions

Motivation and tree-depth. Many efficient graph algorithms are based on decomposing a graph into several simpler pieces on which a problem can be solved efficiently, and then combining the partial solutions to a global solution. In this section we want make use of this divide and conquer principle and cover a graph with (overlapping) pieces such that

(1) the number of pieces is small,

(2) each piece is simple, and

(3) every small subgraph is fully contained in at least one piece.

As a simple algorithmic application of such a covering consider the subgraph isomorphism problem. In this problem we are given two graphs G and H as input, and we are asked to determine whether G contains a subgraph isomorphic to H. In many practical settings the pattern graph H we are looking for is small and in such case a covering as described above is very useful. By the first property, we can iterate through the small number of pieces, by the third property, one of the pieces will contain our pattern graph. By the second property, we can test each piece for containment of H.

We can formulate the covering problem in an equivalent way from the point of view of graph coloring as follows. Color the vertices of a graph G with few colors such that the union of any p color classes induces a simple subgraph. We then interpret the combination of p color classes as a piece in the above formulation. In our case, we will say that a graph is *simple* if it has *bounded tree-depth*.

Definition 4. Let T be a rooted tree. For $u, v \in V(T)$ we write $u \leqslant_T v$ if u lies on the unique shortest path from v to the root of T. The *height* of T is the number of vertices on a longest root-leaf path in T. The *closure* of T, denoted $\mathrm{clos}(T)$, is the graph with vertex set $V(T)$ and $uv \in E(\mathrm{clos}(T))$ if and only if $u <_T v$ or $v <_T u$. A *rooted forest* F is a disjoint union of rooted trees, the height of F is the maximum height among trees $T \in F$. and the closure of F, denoted $\mathrm{clos}(F)$, is the union of all $\mathrm{clos}(T)$ for trees T in F.

Definition 5. Let G be a graph. The *tree-depth* of G, denoted $\mathrm{td}(G)$, is the minimum height of a rooted forest F such that $G \subseteq \mathrm{clos}(F)$.

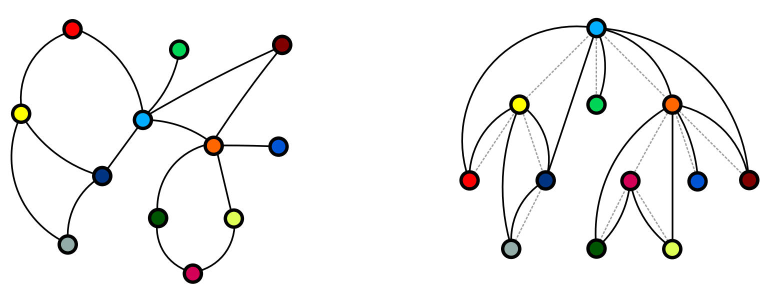

Figure 6: A graph and its tree-depth decomposition of height 4.

A rooted forest F with $G \subseteq \mathrm{clos}(F)$ is often called a *tree-depth decomposition* of G. Note that we may assume w.l.o.g. that $V(F) = V(G)$, for other vertices may be safely removed.

Equivalently, the tree-depth of a graph can be defined using the following recursive formula.

Lemma 14. *For each graph G the following holds:*

$$\mathrm{td}(G) = \begin{cases} 1 & \text{if } |V(G)| = 1 \\ 1 + \min_{v \in V(G)} \mathrm{td}(G - v) & \text{if } G \text{ is connected and } |V(G)| > 1 \\ \max_{i \in \{1,\ldots,k\}} \mathrm{td}(G_i) & \text{if } G_1, \ldots, G_k \text{ are the components of } G. \end{cases}$$

Proof. For $n = 1$ the claim holds vacuously, so consider $n > 1$.

Assume first that G is disconnected, and let $G_1, \ldots, G_k$ be the connected components of G. On one hand, clearly the tree-depth of G is not smaller than the tree-depth of any its connected component, so $\mathrm{td}(G) \geqslant \max_{i \in \{1,\ldots,k\}} \mathrm{td}(G_i)$. On the other hand, if for each connected component G_i we take a rooted forest F_i of height $\mathrm{td}(G_i)$ whose closure contains G_i, then the union of these forests is a rooted forest of height $\max_{i \in \{1,\ldots,k\}} \mathrm{td}(G_i)$ whose closure contains G. This proves that $\mathrm{td}(G) \leqslant \max_{i \in \{1,\ldots,k\}} \mathrm{td}(G_i)$. The two inequalities combined yield $\mathrm{td}(G) = \max_{i \in \{1,\ldots,k\}} \mathrm{td}(G_i)$

Assume now that G is connected. Let T be a rooted forest of height $\mathrm{td}(G)$ whose closure contains G; since G is connected, it is straightforward to see that T must be a rooted tree. Let u be the root of T. Then T with u removed is a rooted forest of height $\mathrm{td}(G) - 1$ whose closure contains the graph $G - u$, witnessing that $\mathrm{td}(G - u) \leqslant \mathrm{td}(G) - 1$. Consequently, $\mathrm{td}(G) \geqslant 1 + \mathrm{td}(G - u) \geqslant 1 + \min_{v \in V(G)} \mathrm{td}(G - v)$. On the other hand, adding one vertex to a graph can increase the tree-depth by at most one, since the new vertex can be always placed as the root of the rooted forest whose closure contains the graph; this shows that $\mathrm{td}(G) \leqslant 1 + \min_{v \in V(G)} \mathrm{td}(G - v)$. By combining the two inequalities, we infer that $\mathrm{td}(G) = 1 + \min_{v \in V(G)} \mathrm{td}(G - v)$. □

Interestingly, tree-depth can be seen as the limit of the sequence of weak coloring numbers for larger and larger radii, as made precise in the following lemma.

Lemma 15. *Let G be an n-vertex graph. Then*

$$\mathrm{wcol}_1(G) \leqslant \mathrm{wcol}_2(G) \leqslant \ldots \leqslant \mathrm{wcol}_n(G) = \mathrm{td}(G).$$

Proof. The only non-trivial assertion is that $\mathrm{wcol}_n(G) = \mathrm{td}(G)$. Clearly, $\mathrm{wcol}_n(G) = \mathrm{wcol}_\infty(G)$ whenever $|V(G)| = n$, so it suffices to prove that $\mathrm{wcol}_\infty(G) = \mathrm{td}(G)$ for all graphs G. To this end, we prove that $\mathrm{wcol}_\infty(G)$ satisfies the same recursive formurla as we gave in Lemma 14 for the tree-depth, namely:

$$\mathrm{wcol}_\infty(G) = \begin{cases} 1 & \text{if } |V(G)| = 1 \\ 1 + \min_{v \in V(G)} \mathrm{wcol}_\infty(G - v) & \text{if } G \text{ is connected and } |V(G)| > 1 \\ \max_{i \in \{1,\ldots,k\}} \mathrm{wcol}_\infty(G_i) & \text{if } G_1, \ldots, G_k \text{ are the components of } G. \end{cases}$$

If the above formula indeed holds, then the lemma will follow by a straightforward induction on the number of vertices of G. The formula holds trivially for $|V(G)| = 1$, so assume $|V(G)| > 1$.

First assume that G is disconnected and let $G_1, \ldots, G_k$ be the connected components of G. As in the proof of Lemma 14, we clearly have that $\mathrm{wcol}_\infty(G) \geqslant \mathrm{wcol}_\infty(G_i)$ for each connected component G_i, implying $\mathrm{wcol}_\infty(G) \geqslant \max_{i \in \{1,\ldots,k\}} \mathrm{wcol}_\infty(G_i)$, whereas the reverse inequality follows from observing that the concatenation of optimum vertex orderings for $G_1, \ldots, G_k$ is a vertex ordering for G with the weak ∞-coloring number equal to $\max_{i \in \{1,\ldots,k\}} \mathrm{wcol}_\infty(G_i)$.

Second, assume that G is connected. Let σ be a vertex ordering of G with the optimum weak ∞-coloring number, and let u be the vertex of G that is the smallest in σ. Since G is connected, we have that $u \in \mathrm{WReach}_\infty[G,\sigma,v]$ for all $v \in V(G)$. This implies that removing u from σ yields a vertex ordering of $G-u$ with weak ∞-coloring number at most $\mathrm{wcol}_\infty(G)-1$, implying that $\mathrm{wcol}_\infty(G) \geqslant 1+\mathrm{wcol}_\infty(G-u) \geqslant 1+\min_{v\in V(G)} \mathrm{wcol}_\infty(G-v)$. On the other hand, adding a vertex to a graph can increase its weak ∞-coloring number by at most 1, for the new vertex can be always placed as the smallest in the ordering; this implies that $\mathrm{wcol}_\infty(G) \leqslant 1+\min_{v\in V(G)} \mathrm{wcol}_\infty(G-v)$. The two inequalities prove that indeed $\mathrm{wcol}_\infty(G) = 1+\min_{v\in V(G)} \mathrm{wcol}_\infty(G-v)$. □

Low tree-depth colorings. We now come to the definition of *low tree-depth decompositions* or *low tree-depth colorings.*

Definition 6. Let G be a graph and let $r \in \mathbb{N}$. An *r-tree-depth coloring* of G is a coloring of vertices of G with some set of colors such that any $r' \leqslant r$ color classes induce a subgraph with tree-depth at most r'.

Thus, in an r-tree-depth coloring, every color class must be an independent set, every pair of color classes must induce a forest of stars, and so on up to depth r. We will now prove that the weak coloring numbers and low tree-depth colorings are strongly related, and in particular classes of bounded expansion can be characterized as those that admit r-tree-depth colorings with a bounded number of colors, for every $r \in \mathbb{N}$.

Lemma 16. *Let G be a graph and let $r \in \mathbb{N}$. If $\mathrm{wcol}_{2^{r-2}}(G) \leqslant m$, then the vertices of G can be colored with m colors so that any for connected subgraph $H \subseteq G$, either some color appears exactly once in H or H receives at least r distinct colors.*

Proof. Let σ be an ordering of $V(G)$ with $\mathrm{wcol}_{2^{r-2}}(G,\sigma) \leqslant m$. Color the vertices greedily with m colors, from left to right along the order σ, such that the color assigned to a vertex v is distinct from all colors assigned to vertices weakly (2^{r-2})-reachable from v. We claim that this coloring, call it λ, satisfies the desired properties.

Let H be a connected subgraph of G and let v be the minimum vertex of H with respect to σ. If the color $\lambda(v)$ appears exactly once in H, then we are done.

Hence assume that $\lambda(v)$ occurs more than once in H. We shall prove that H receives at least r different colors. To this end, we will find paths $P_1,\ldots,P_{r-1}$ such that

$$H \supseteq P_1 \supseteq P_2 \supseteq \ldots \supseteq P_{r-1}$$

and vertices $u_0,u_1,\ldots,u_{r-2}$ with

$$u_0 \in V(H)-V(P_1) \qquad \text{and} \qquad u_i \in V(P_i)-V(P_{i+1}) \text{ for } 1 \leqslant i \leqslant r-2$$

such that the color $\lambda(u_i)$ does not appear in P_j for all $0 \leqslant i < j \leqslant r-1$. Furthermore, we will guarantee that $|V(P_i)| \geqslant 2^{r-i-1}$ for all $1 \leqslant i \leqslant r-1$, which in particular implies $|V(P_{r-1})| \geqslant 1$. Hence, the colors $\lambda(u_i)$ for $0 \leqslant i \leqslant r-2$ are all distinct and we can find one additional vertex $u_{r-1} \in V(P_{r-1})$ whose color is distinct from all other colors. This gives us r distinct colors in total.

Let $u \neq v$ be a vertex of H with $\lambda(u) = \lambda(v)$ and let $P = v, v_1, \ldots, v_q = u$ be any path in H connecting v and u; such path exists since H is connected. We must have $q > 2^{r-2}$, for otherwise v would be weakly (2^{r-2})-reachable from u and we would have $\lambda(v) \neq \lambda(u)$. Let $u_0 \coloneqq v$ and let $P_1 \coloneqq v_1,\ldots,v_{2^{r-2}}$. Clearly P_1 has 2^{r-2} vertices and with the same argument as above, no vertex of P_1 has color $\lambda(u_0)$, as u_0 is weakly (2^{r-2})-reachable from every vertex of P_1.

If the paths $P_1, \ldots, P_i$ have been constructed and satisfy the above conditions, we can repeat the above argument to find u_i and P_{i+1} with the desired properties. Simply let u_i be the vertex which is the smallest with respect to σ on P_i and argue as above that its color under λ is unique on P_i, for it is weakly (2^{r-2})-reachable from every vertex of P_i. Now let P_{i+1} be the larger of the two subpaths into which the removal of u_i breaks P_i. Since P_i contained at least 2^{r-i-1} vertices, it follows that P_{i+1} contains at least 2^{r-i-2} vertices. □

As we will see in a moment, the properties of the coloring yielded by Lemma 16 in fact guarantee that it is a low tree-depth coloring. We give a special name to such colorings.

Definition 7. An *r-centered coloring* of a graph G is a coloring of vertices of G such that for any connected subgraph $H \subseteq G$, either some color appears exactly once in H or H receives more than r different colors.

Lemma 17. *Any r-centered coloring of a graph is also an r-tree-depth coloring.*

Proof. Let λ be an r-centered coloring of G. Assume, for the sake of contradiction, that there is a subgraph $G' \subseteq G$ with $\mathrm{td}(G') = k \leqslant r$ which receives less than k colors. Choose G' to be minimal with this property. Then G' is connected. As G' receives less than $k \leqslant r$ colors and the coloring λ is r-centered, there is one color which occurs exactly once in G', say this color is given to vertex v. Then $\mathrm{td}(G' - v) \geqslant k - 1$ and $G' - v$ receives less than $k - 1$ colors. Hence G' is not minimal with the considered property, contradicting our assumption. □

Conversely, supposing that a graph G admits a low tree-depth coloring, we can bound the density of depth-r topological minors in G, and hence by Theorem 9 also the weak coloring numbers of G are bounded. We first need the following simple claim.

Lemma 18. *For any graph G, it holds that $\widetilde{\nabla}_\infty(G) \leqslant \mathrm{td}(G) - 1$.*

Proof. It suffices to show that for every topological minor H of G contains a vertex of degree at most $\mathrm{td}(G) - 1$. Let F be a rooted forest of height $\mathrm{td}(G)$ whose closure contains G. Let ϕ be a minor model of H in G. Let u be a vertex of H such that $\phi(u)$ is at the largest depth in F among the vertices of $\phi(V(H))$. Hence, for every edge $uv \in E(H)$ the vertex $\phi(v)$ is not a descendant of $\phi(u)$ in F, so the path $\phi(uv)$ has to contain a strict ancestor of $\phi(u)$ in F. All these strict ancestors have to be pairwise different, and the number of such ancestors is at most $\mathrm{td}(G) - 1$, hence the number of neighbors v of u in H is at most $\mathrm{td}(G) - 1$. □

Lemma 19. *Let G be a graph and let $r \in \mathbb{N}$. Assume that G admits a $(2r+1)$-tree-depth coloring with m colors. Then $\widetilde{\nabla}_r(G) \leqslant 2r \cdot \binom{m}{2r+1}$.*

Proof. Fix a coloring $\lambda \colon V(G) \to \{1, \ldots, m\}$ such that the union of any $i \leqslant 2r+1$ color classes in λ induces a graph of tree-depth at most i. Let H be a depth-r topological minor of G, and let ϕ be a depth-r topological minor model of H in G. We have to show that $|E(H)| \leqslant 2r \cdot \binom{m}{2r+1} \cdot |V(H)|$.

Let $\mathcal{M}$ be the set of all subsets of $\{1, \ldots, m\}$ of size $2r+1$. Every edge $e \in E(H)$ corresponds to a path $\phi(e)$ of length at most $2r+1$ in G. Hence, we can partition $E(H)$ into $\binom{m}{2r+1}$ sets $\{E_I\}_{I \in \mathcal{M}}$ such that an edge e may belong to E_I only if I contains all colors that occur on $\phi(e)$. For each $I \in \mathcal{M}$, let H_I be the subgraph of H consisting of edges of E_I and vertices incident to them. It follows from the assumed property of E_I that the images of all edges of H_I under ϕ are contained in a subgraph of G induced by $\lambda^{-1}(I)$; this subgraph has tree-depth at most $2r+1$ by the assumption that λ is a $(2r+1)$-tree-depth coloring. Hence, H_I is a topological minor of a graph of tree-depth at most $2r+1$, so by Lemma 18 we have

$$|E_I| \leqslant 2r \cdot |V(H_I)| \leqslant 2r \cdot |V(H)|.$$

By summing the inequalities as above for all $I \in \mathcal{M}$, we conclude that $|E(H)| \leqslant 2r\binom{m}{2r+1} \cdot |V(H)|$, as requested. □

Lemmas 16, 17, and 19, together with Theorem 9 and the relations between density of shallow minors and of shallow topological minors, yield the following.

Theorem 20. *Let $\mathcal{C}$ be a class of graphs. Then the following conditions are equivalent.*

(a) $\mathcal{C}$ has bounded expansion.

(b) There is a function $M\colon \mathbb{N} \to \mathbb{N}$ such that for each $r \in \mathbb{N}$, every graph $G \in \mathcal{C}$ admits an r-tree-depth coloring with $M(r)$ colors.

Algorithmic applications. Let us dwell a bit on the algorithmic aspects of Theorem 20. Fix a class $\mathcal{C}$ of bounded expansion and a graph $G \in \mathcal{C}$. By Theorem 6, we can compute in time $\mathcal{O}(n^4)$ an vertex ordering of G with r-approximate r-admissibility, for any constant r we choose, using the fact that on bounded expansion classes we have $\mathcal{O}(n^3 m) = \mathcal{O}(n^4)$. By Corollary 4, this ordering has also the weak r-coloring number bounded by a constant. Now, the construction of Lemma 16 can be easily made algorithmic, and by Lemma 17 the obtained coloring is a p-treedepth coloring, assuming we chose $r = 2^{p-2}$. This yields the following.

Corollary 21. *For a fixed $p \in \mathbb{N}$ and a class of bounded expansion $\mathcal{C}$, given a graph $G \in \mathcal{C}$ one can compute a p-treedepth coloring of G with $M(p)$ colors in time $\mathcal{O}(n^4)$.*

We can hence efficiently solve the subgraph isomorphism problem on classes of bounded expansion, as hinted at in the introduction of this section. We first note that we can solve it efficiently on graphs of bounded tree-depth.

For this, it is essential that we are able to approximate the tree-depth of a graph G and compute a low height forest F such that $G \subseteq \mathrm{clos}(G)$. It turns out that there is a very simple way to do it, provided we are happy with obtaining height exponential in the optimum. Namely, any depth-first search forest of a graph G is a valid tree-depth decomposition of G of height at most $2^{\mathrm{td}(G)}$, as we show in the next lemmas.

Lemma 22. *The tree-depth of an n-vertex path is equal to $\lceil \log_2(n+1) \rceil$.*

Proof. It suffices to prove that whenever $2^k \leqslant n \leqslant 2^{k+1}-1$, the tree-depth of P_n is equal to $k+1$. We proceed by induction on k. For $k=0$ the claim holds trivially. Let P_n be the path on n vertices, where $2^k \leqslant n \leqslant 2^{k+1}-1$.

We first show that $\mathrm{td}(P_n) \leqslant k+1$. Let u be a middle vertex on P_n, that is, one whose removal splits P_n into two subpaths on at most $\lfloor n/2 \rfloor < 2^k$ vertices each. By induction assumption, for each of these subpaths we can find a tree-depth decomposition of height at most k, and these can be combined into a tree-depth decomposition of P_n of height $k+1$ by taking their union and adding u as the root.

We now show that $\text{td}(P_n) \geqslant k+1$. Take any tree-depth decomposition T of P_n; T is a rooted tree since P_n is connected. Let u be the root of T. Then each of the two subpaths of $P_n - u$ is placed entirely in one subtree rooted at a child of u in T. Since $n \geqslant 2^k$, one of these subpaths has at least 2^{k-1} vertices, so by the induction assumption its tree-depth is at least k. Then the corresponding subtree of T rooted at a child of u has height at least k, implying that T has height at least $k+1$. □

Lemma 23. *Let G be a graph. Then every rooted forest F obtained by running a depth-first search in each connected component of G satisfies $G \subseteq \text{clos}(F)$ and $\text{height}(F) < 2^{\text{td}(G)}$.*

Proof. Let F be such a rooted forest. It is straightforward to see that $G \subseteq \text{clos}(F)$. Indeed, if there was an edge $uv \in E(G)$ such that u and v were not bound by the ancestor-descendant relation in F, then provided u was visited earlier by the DFS than v, the edge uv would be used by the DFS to access v from u, so v should have been a descendant of u. To see that $\text{height}(F) \leqslant 2^{\text{td}(G)}$, observe that if $d \coloneqq \text{height}(F)$, then G contains a path on d vertices. By Lemma 22 we infer that the tree-depth of this path, and consequently also the tree-depth of G, is at least $\lceil \log_2(d+1) \rceil$. Since $\text{td}(G) \geqslant \lceil \log_2(d+1) \rceil > \log_2 d$, it follows that $d < 2^{\text{td}(G)}$. □

With an approximate decomposition at hand, we can solve the subgraph isomorphism problem on graphs of bounded tree-depth.

Lemma 24. *Let G, H be graphs and assume that G has tree-depth at most k. Then we can decide in time $f(k, H) \cdot |V(G)|$ whether G contains a subgraph isomorphic to H, for some computable function f.*

Proof sketch. By Lemma 23, in linear time we can compute a tree-depth decomposition of G of depth at most $d \coloneqq 2^k$. On this decomposition one can employ a simple, though a bit tedious dynamic programming algorithm with running time $d^{\mathcal{O}(|V(H)|)} \cdot n$. Details will be given during the tutorials. □

We can now solve the subgraph isomorphism problem on classes of bounded expansion.

Theorem 25. *Let $\mathcal{C}$ be a class of bounded expansion and let H be a graph. Then the subgraph isomorphism problem with pattern graph H can be decided on $\mathcal{C}$ in time $f(H) \cdot n^4$.*

Proof. Let $h = |V(H)|$. According to Corollary 21, we can compute in time $f(h) \cdot n^4$ an h-tree-depth coloring for an n-vertex input graph $G \in \mathcal{C}$ with $M(h)$ colors, for some functions M and f. Now we iterate through all combinations of at most h colors and test for each such combination I, whether H is isomorphic to a subgraph of G_I, where G_I is the subgraph of G induced by the union of colors of I. As H has only h vertices, we can give a positive answer to the subgraph isomorphism problem if and only if we find a subgraph G_I containing H. As each G_I has tree-depth at most h, according to Lemma 24 we can test this in time $g(h) \cdot n$ for some function g. Therefore, by iterating through all $\binom{M(h)}{h}$ combinations I of h colors, we obtain the overall running time of $\binom{M(h)}{h} \cdot g(h) \cdot n$. □

The bottleneck of the computation in Theorem 25 is the computation of an ordering with bounded weak coloring number. As we mentioned in Theorem 5, on every bounded expansion class there is a linear time algorithm for computing r-admissibility exactly, which hence can be used as a linear time algorithm for approximating the weak coloring number. Using this result, plus a number of technical checks of implementation details, one can improve the running time of the algorithm of Theorem 25 to linear $f(H) \cdot n$ for some function f.

A careful reader has probably observed that in fact, the usage of the approximation algorithm for tree-depth in the proof of Theorem 25 was actually not necessary. This is because Corollary 21 yields not only a h-tree-depth coloring, but actually a h-centered coloring, and in an h-centered coloring it is straightforward to construct tree-depth decomposition for every combination of h colors, essentially by simulating the proof of Lemma 17. However, the ability of approximating tree-depth shows that low tree-depth colorings can be used as a black-box, without understanding the inner workings of the proof of their existence.

6 Digression on logic

It turns out that the subgraph isomorphism problem we solved in the previous section is a special case of a much more general theorem. We can in fact evaluate every fixed first-order formula on a class of bounded expansion in linear time.

First-order formulas over the vocabulary of graphs are formed from atomic formulas $x = y$ and $E(x, y)$, where x, y are variables (we assume that we have an infinite supply of variables), using the usual Boolean connectives $\neg$ (negation), $\wedge$ (conjunction), and $\vee$ (disjunction) and existential and universal quantification $\exists x, \forall x$, respectively. The set of all first-order formulas is denoted by FO. The free variables of a formula are those not in the scope of a quantifier, and we write $\varphi(x_1, \ldots, x_k)$ to indicate that the free variables of the formula φ are among $x_1, \ldots, x_k$. A *sentence* is a formula without free variables. The *quantifier rank* $\mathrm{qr}(\varphi)$ of a formula φ is the nesting depth of quantifiers in φ, defined recursively in the obvious way. A formula without any quantifiers is called *quantifier-free*. To define the semantics, we inductively define a satisfaction relation $\models$, where for a graph G, a formula $\varphi(x_1, \ldots, x_k)$, and vertices $v_1, \ldots, v_k \in V(G)$,

$$G \models \varphi(v_1, \ldots, v_k)$$

means that G satisfies φ if the free variables $x_1, \ldots, x_k$ are interpreted by $v_1, \ldots, v_k$, respectively. We have $G \models x_1 = x_2$ if and only if $v_1 = v_2$ and $G \models E(x_1, x_2)$ if and only if $v_1 v_2 \in E(G)$. Similarly, the meaning of the Boolean connectives and the quantifiers is exactly as expected.

Let us give some examples. For a fixed graph H with vertex set $\{v_1, \ldots, v_h\}$ consider the formula

$$\varphi_H = \exists x_1 \ldots \exists x_h \left(\bigwedge_{i<j} x_i \neq x_j \wedge \bigwedge_{i<j\,:\,v_i v_j \in E(G)} E(x_i, x_j) \right).$$

Then for every graph G we have

$$G \models \varphi_H \qquad \text{if and only if} \qquad H \text{ is isomorphic to a subgraph of } G.$$

As another example, consider the formula

$$\psi(x_1, \ldots, x_k) = \forall y \left(\bigvee_{1 \leqslant i \leqslant k} (y = x_i) \vee E(y, x_i) \right).$$

Then for every graph G and vertices $v_1, \ldots, v_k \in V(G)$ we have $G \models \psi(v_1, \ldots, v_k)$ if and only if $\{v_1, \ldots, v_k\}$ is a dominating set of G. Finally, for an integer r, consider the formula $\delta_r(x, y)$ defined as follows:

$$\delta_r(x, y) = \exists z_0 \ldots \exists z_r \, (x = z_0) \wedge (y = z_r) \wedge \bigwedge_{i=1}^{r} ((z_i = z_{i-1}) \vee E(z_i, z_{i-1}))\,.$$

Then for vertices $u, v \in V(G)$, we have $G \models \delta_r(u, v)$ if and only if the distance between u and v in G is at most r.

Hence, first-order logic can express the subgraph isomorphism problem, dominating set problem, and many more important graph problems. It appears that checking whether an FO sentence holds in any bounded expansion class can be done in linear time, and an almost linear time can be claimed on nowhere dense classes.

Theorem 26. *Let $\mathcal{C}$ be a class of bounded expansion. There exists a function f such that given a first-order formula φ and an n-vertex graph $G \in \mathcal{C}$, it can be tested whether $G \models \varphi$ in time $f(|\varphi|) \cdot n$.*

Theorem 27. *Let $\mathcal{C}$ be a nowhere dense class. For every $\varepsilon > 0$ there exists a function f such that given a first-order formula φ and an n-vertex graph $G \in \mathcal{C}$, it can be tested whether $G \models \varphi$ in time $f(|\varphi|) \cdot n^{1+\varepsilon}$.*

Theorems 26 and 27 are difficult and require an extensive introduction to logic in computer science. Therefore, we will not cover them in this course. However, let us mention that at least for Theorem 26, low tree-depth colorings or orderings with bounded weak coloring numbers can be used as main combinatorial tools.

7 Neighborhood complexity

The next topic of our structural investigations is *neighborhood complexity*. The setting is as follows. Suppose we have fixed some radius $r \in \mathbb{N}$, and we investigate a graph G and a subset of its vertices $A \subseteq V(G)$. Suppose further that what interests us is how different vertices of the graph "interact" with A, where interaction is some "local" relation. The simplest example, on which we will focus, is just the distance relation: for a vertex $u \in V(G)$, the vertices of A with which u interacts are simply those that are at distance at most r from u; in other words, we investigate the *r-neighborhood of u in A* defined as $N_G^r[u] \cap A$. The question is: how many different interactions (r-neighborhoods) can there be? It turns out that in classes of bounded expansion this number is always *linear* in the size of A, as explained in the following theorem, on which we will focus now.

Theorem 28. *Let $\mathcal{C}$ be a class of bounded expansion and let $r \in \mathbb{N}$. There exists a constant c, depending only on $\mathcal{C}$ and r, such that for every $G \in \mathcal{C}$ and nonempty $A \subseteq V(G)$, we have*

$$|\{N_G^r[u] \cap A \colon u \in V(G)\}| \leqslant c \cdot |A|.$$

Actually, we will prove a stronger result: not only the number of different r-neighborhoods is linear, but even a number of *r-distance profiles* is linear, where an r-distance profile is r-neighborhood enriched with information on the actual values of distances not greater than r.

Definition 8. For $r \in \mathbb{N}$, a graph G, a vertex subset $A \subseteq V(G)$, and a vertex $u \in V(G)$, the *r-distance profile* of u on A is the function $\pi_r[u, A]\colon A \to \{0, 1, \ldots, r, \infty\}$ defined as follows: for $a \in A$, if $\mathrm{dist}_G(u, a) \leqslant r$ then $\pi_r[u, A](a) = \mathrm{dist}_G(u, a)$, and otherwise $\pi_r[u, A](a) = \infty$. A function $f\colon A \to \{0, 1, \ldots, r, \infty\}$ is *realized* as an r-distance profile on A if there exists $u \in V(G)$ with $f = \pi_r[u, A]$.

Observe that two vertices with the same r-distances profiles on A have the same r-neighborhood in A, hence to prove Theorem 28 it suffices to prove the following.

Theorem 29. *Let $\mathcal{C}$ be a class of bounded expansion and let $r \in \mathbb{N}$. There exists a constant c, depending only on $\mathcal{C}$ and r, such that for every $G \in \mathcal{C}$ and nonempty $A \subseteq V(G)$, the number of different functions from A to $\{0, 1, \ldots, r, \infty\}$ realized as r-distance profiles on A is at most $c \cdot |A|$.*

Proof. Denote $d \coloneqq \mathrm{wcol}_{2r}(\mathcal{C})$; since $\mathcal{C}$ has bounded expansion, d is a constant depending only on r and $\mathcal{C}$. We will give a proof that the number of different functions realized as r-projection profiles on A is bounded by

$$1 + d \cdot 2^{d-1} \cdot (r+2)^d \cdot |A|,$$

hence setting $c \coloneqq 1 + d \cdot 2^{d-1} \cdot (r+2)^d$ will suffice.

Since $G \in \mathcal{C}$, there is some vertex ordering σ of G with $\mathrm{wcol}_{2r}(G, \sigma) \leqslant d$; we fix such an ordering for the rest of the proof. For brevity, for $u \in V(G)$ we write $\mathrm{WReach}_r[u]$ instead of $\mathrm{WReach}_r[G, \sigma, u]$.

Let $B \coloneqq \bigcup_{a \in A} \mathrm{WReach}_r[a]$. Obviously $B \supseteq A$ and $|B| \leqslant d|A|$, because $|\mathrm{WReach}_r[a]| \leqslant \mathrm{wcol}_{2r}(G, \sigma) \leqslant d$ for each $a \in A$. For $u \in V(G)$, we define the *local separator* of u as

$$X[u] \coloneqq \mathrm{WReach}_r[u] \cap B.$$

The name is justified by the fact that $X[u]$ is a separator for paths of length at most r from u to A in the following sense.

Claim 3. *For each vertex $u \in V(G)$, every path of length at most r connecting u with a vertex of A contains a vertex of $X[u]$.*

Proof. Take any such path P and let x be the smallest vertex on P in the order σ. Then the prefix of P up to x and the suffix of P from x onward witness that $x \in \mathrm{WReach}_r[u] \cap B$. ⌟

We next show that if two vertices have the same local separator and the same r-distance profile on it, then they actually have the same r-distance profile on A. This corroborates the intuition that all the "information flow" between a vertex and A has to pass through its local separator.

Claim 4. *For every pair of vertices $u, v \in V(G)$, if $X[u] = X[v] = X$ for some $X \subseteq B$ and $\pi_r[u, X] = \pi_r[v, X]$, then also $\pi_r[u, A] = \pi_r[v, A]$.*

Proof. By symmetry, it suffices to show that if for some $a \in A$ and $q \leqslant r$ we have $\mathrm{dist}_G(u, a) \leqslant q$, then also $\mathrm{dist}_G(v, a) \leqslant q$. Let P be a path of length at most q that connects u and a. Since $q \leqslant r$, by Claim 3 we have that P contains a vertex of X, say x. In particular P witnesses that $\mathrm{dist}_G(u, x) + \mathrm{dist}_G(x, a) \leqslant q \leqslant r$. Since $\pi_r[u, X] = \pi_r[v, X]$ and $\mathrm{dist}_G(u, x) \leqslant r$, we have that $\mathrm{dist}_G(u, x) = \mathrm{dist}_G(v, x)$. Consequently, by the triangle inequality we have

$$\mathrm{dist}_G(v, a) \leqslant \mathrm{dist}_G(v, x) + \mathrm{dist}_G(x, a) = \mathrm{dist}_G(u, x) + \mathrm{dist}_G(x, a) \leqslant q,$$

and we are done. ⌟

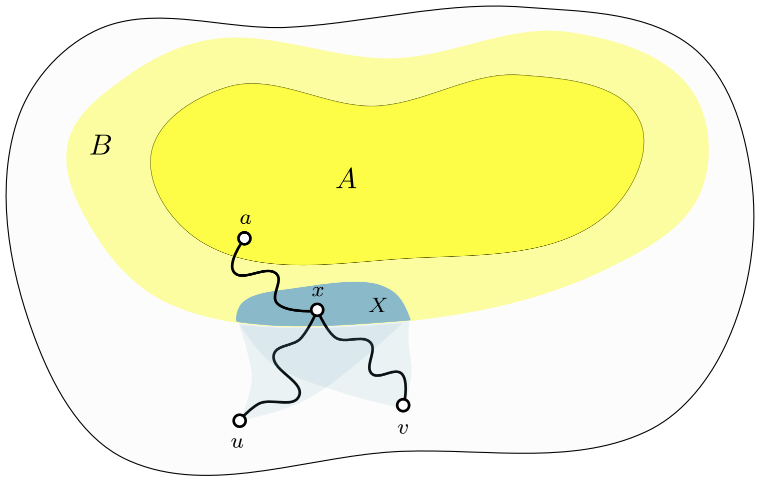

Figure 7: Situation in the proof of Claim 4.

Let

$$\mathcal{X} \coloneqq \{X[u] \colon u \in V(G)\} - \{\emptyset\}.$$

In other words, $\mathcal{X}$ comprises all nonempty subsets of B that are realized as a local separator for some vertex u. Observe that since for each $u \in V(G)$ we have $X[u] \subseteq \mathrm{WReach}_r[u]$ and $|\mathrm{WReach}_r[u]| \leqslant \mathrm{wcol}_{2r}(G,\sigma) \leqslant d$, each set in $\mathcal{X}$ has size at most d. Thus, for a given $X \in \mathcal{X}$ the number of different r-distance profiles on X is bounded by the total number of functions from X to $\{0,1,\ldots,r,\infty\}$, which in turn is upper bounded by $(r+2)^d$. Therefore, it suffices to prove that

$$|\mathcal{X}| \leqslant d \cdot 2^{d-1} \cdot |A|. \tag{7.1}$$

Indeed, then on each local separator from $\mathcal{X}$ we will have at most $(r+2)^d$ different r-neighborhood profiles, yielding, by Claim 4, at most $d \cdot 2^{d-1} \cdot (r+2)^d \cdot |A|$ different r-neighborhood profiles on A. The additional $+1$ summand is because we also need to take into consideration vertices u with $X[u] = \emptyset$; again by Claim 4, all those vertices have the same profile on A (it is not hard to see that this profile maps every vertex of A to ∞).

Hence we are left with proving (7.1). For $X \in \mathcal{X}$ let us define

$$\varphi(X) \coloneqq \text{the largest vertex of } X \text{ in the ordering } \sigma.$$

Thus, φ is a function from $\mathcal{X}$ to B. Since $|B| \leqslant d|A|$, to show (7.1) it suffices to prove the following.

Claim 5. *For each $b \in B$ we have $|\varphi^{-1}(b)| \leqslant 2^{d-1}$.*

Proof. It suffices to show that for each $X \in \varphi^{-1}(b)$ we have $X \subseteq \mathrm{WReach}_{2r}[b]$. Indeed, since $|\mathrm{WReach}_{2r}[b]| \leqslant d$, there are at most 2^{d-1} different subsets of $\mathrm{WReach}_{2r}[b]$ containing b, and by the assertion above these will be the only candidates for sets from $\varphi^{-1}(b)$.

Let then $X \in \mathcal{X}$ be such that $\varphi(X) = b$, that is, b is the largest element of X in σ. Since $X \in \mathcal{X}$, there is a vertex $u \in V(G)$ such that $X = X[u]$. Take any $x \in X$. As $X \subseteq \mathrm{WReach}_r[u]$, there is a path P of length at most r from u to x such that x is the smallest vertex traversed by P in the ordering σ. Similarly, since $b \in X \subseteq \mathrm{WReach}_r[u]$, there is a path Q of length at most r from u to b such that b is the smallest vertex traversed by P in the ordering σ. By the choice of b we

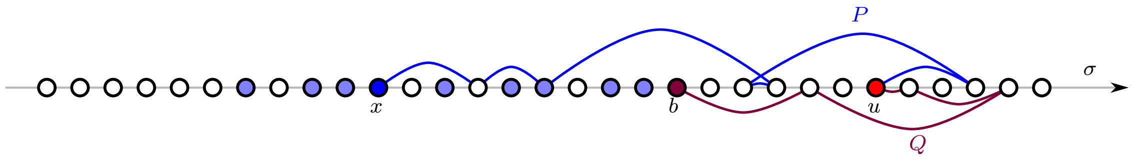

Figure 8: Situation in the proof of Claim 5. The vertices of X are depicted in blue. The concatenation of paths P and Q is a walk of length at most $2r$ that witnesses that $x \in \mathrm{WReach}_{2r}[b]$.

have that $x \leqslant_\sigma b$, hence all vertices traversed by P or Q are not smaller than x in σ. We conclude that the concatenation of P and Q witnesses that $x \in \mathrm{WReach}_{2r}[b]$. As x was chosen arbitrarily, this implies that $X \subseteq \mathrm{WReach}_{2r}[b]$ and finishes the proof. ⌟

As we discussed, Claim 5 implies (7.1), which in turn implies the statement of the theorem. □