Chapter 1:
Measuring sparsity

Compilation date: January 24, 2020

1 Motivating examples and concepts

Before we start delving into the theory of structural sparsity, we present a number of classic concepts in graph theory, on which many of abstract notion will built later on. These concepts include both classes of sparse graphs and various decomposition notions.

1.1 Planar graphs

We start our story with probably the most important class of starse graphs: *planar graphs*. Since we do not want to put an emphasis on the topological aspects of graph embeddings (and only slide over embeddings into surfaces of higher genera), we will rely here on the intuitive understanding of an embedding. Readers interested in more intricate aspects of graph embeddings are referred to an excellent monograph of Mohar and Thomassen.

A *plane embedding* of a graph G is a mapping that assigns to every vertex of G a distinct point of the two dimensional sphere and to every edge of G a Jordan curve without self-intersections that connects the images of the endpoints of the edge. Distinct edges share no points except for the necessary case: when two edges are incident with the same vertex, the images of the curves meet at their endpoints at the image of the vertex in question. A *plane graph* is a graph G together with its plane embedding. A graph G is *planar* if it admits a plane embedding.

A *face* of an embedding is a connected component of the sphere minus the image of the graph. An embedding is *cellular* if every face is homomorphic to a disc. Note that connected graphs will have only cellular plane embeddings, while disconnected ones will not admit a cellular embedding into a sphere. For this reason, we henceforth consider only connected graphs and thus only cellular plane embeddings.

An astute reader my wonder why we define an embedding as a one into a sphere, not into the plane $\mathbb{R}^2$. The only difference is that an embedding into $\mathbb{R}^2$ distinguishes one infinite face as the *outerface*, which we prefer not to do unless it is essential for the argument.

In this lecture we will be interested what can be proven about planar graphs *using only arguments based on density*. That is, we will try to refrain from using any topological-based arguments. The main "sparsity" tool for planar graphs is the well-known Euler's formula.

Theorem 1.1 (Euler's formula)**.** *If G is a connected plane graph with n vertices, m edges, and f faces, then*

$$n + f - m = 2.$$

Proof. We prove the statement by induction on the number of edges of G. In the base case, G is a tree: $f = 1$ and $m = n - 1$, yielding the statement.

In the inductive case, G is not a tree, and hence contains a cycle C. Pick an edge e on C. The existence of C implies that (i) e is incident with two distinct faces, and (ii) $G - e$ is still connected. Hence, the induced embedding of $G - e$ has one less edge than G and one less face than G. The Euler's formula for G follows from the inductive hypothesis for $G - e$. □

Corollary 1.2. *A connected planar graph with n vertices has less than $3n$ edges and has a vertex of degree at most 5.*

Proof. We present here a slightly longer proof using the method of *discharging*. Discharging is nothing more than a double-counting argument using the Euler's formula. However, in many cases it turns to be the most clean way of expressing double-counting arguments on planar graphs. Fix a plane embedding of an n-vertex connected graph G.

Setup. Every vertex receives a charge of $+6$, every face receives a charge of $+6$, and every edge receives a charge of -6. Euler's formula implies that the total sum of all assigned charges is $+12$.

Discharging rules. We move the charge according to the following two rules.

discharging a face Every face splits its charge evenly between all incident edges. (If a single edge appears on a face twice, then it receives a double share of the charge.) Note that every face gets rid of all its charge.

discharging a vertex Every vertex sends a charge of $+1$ to every incident edge. Note that the final charge of a vertex v is $6 - \mathrm{d}_G(v)$.

Final charge analysis. The crucial observation is that every edge ends up with a nonpositive charge: it starts with a charge of -6, receives twice $+1$ from both endpoints, and twice at most $+2$ from both incident faces, because a face splits a charge of $+6$ between at least three incident edges. With no final charge on faces, the initial charge of $+12$ needs to end up somewhere, and the only place for the positive charge are vertices.

First, this implies an existence of a vertex with positive charge, and thus of degree less than 3. Second, we obtain that

$$\sum_{v \in V(G)} 6 - \mathrm{d}_G(v) \geq 12.$$

Since the sum of degrees of vertices equals twice the number of edges of the graph (every edge is counted exactly twice), the above is equvalent to

$$6n - 2|E(G)| \geq 12,$$

which implies $|E(G)| < 3n$, as desired. □

Corollary 1.2 motivates the following notions:

minimum degree of a graph G is $\delta(G) = \min_{v \in V(G)} \mathrm{d}_G(v)$.

maximum degree of a graph G is $\Delta(G) = \max_{v \in V(G)} \mathrm{d}_G(v)$.

average degree of a graph G is

$$\mathsf{avgdeg}(G) = \frac{\sum_{v \in V(G)} \mathrm{d}_G(v)}{|V(G)|} = 2 \cdot \frac{|E(G)|}{|V(G)|}.$$

Thus, Corollary 1.2 can be now expressed in the new language as:

Corollary 1.3. *A connected planar graph has average degree less than* 6 *and minimum degree at most* 5*.*

From Corollary 1.3 we can deduce the following statement. Recall that a *k-coloring* of a graph G is a function $c : V(G) \to \{1, 2, \ldots, k\}$ such that $c(u) \neq c(v)$ for every $uv \in E(G)$, a graph G is *k-colorable* if it admits a k-coloring, and the *chromatic number* of a graph G, denoted $\chi(G)$, is the minimum k such that G is k-colorable.

Theorem 1.4 (6-color theorem)**.** *Every planar graph is* 6*-colorable.*

Proof. We prove the statement by induction on $|V(G)|$. In the base case, an empty graph is clearly 6-colorable. Otherwise, by Corollary 1.3 there exists $v \in V(G)$ of degree at most 5. By the inductive hypothesis, there exists a 6-coloring c of $G - v$. Since the degree of v is at most 5, there exists a color $i \in \{1, 2, \ldots, 6\}$ that is not assigned by c to any of the neighbors of v. Thus, extending c with $c(v) = i$ we obtain a 6-coloring of G. □

At first glance, it seems that in the proof of Theorem 1.4 we have only used the fact that the minimum degree of G is at most 5 (or average degree less than 6). However, due to the use of induction, we have in fact relied on the fact that *every subgraph of G has minimum degree at most* 5.

To see the difference, consider an example of a big n-vertex clique with huge — say n^3 — number of degree-1 vertices attached to this clique. Such a graph has average degree $1 + \mathcal{O}(n^{-1})$, but is clearly not k-colorable for any $k < n$. This is because it contains a dense subgraph, the n-vertex clique, and the additional pendant vertices only disguises it in the average degree count.

1.2 Degeneracy

Hence, it makes sense to introduce notions that are "closed under taking subgraphs" analogs of minimum and average degree. Here, $H \subseteq G$ denotes that H is a subgraph of G.

Definition 1.5. For a graph G, its

maximum average degree is $\mathsf{mad}(G) = \max_{H \subseteq G} \mathsf{avgdeg}(G)$.

degeneracy is $\mathsf{deg}(G) = \max_{H \subseteq G} \delta(G)$.

In other words, G is *of degeneracy at most d* if every subgraph of G contains a vertex of degree at most d.

Clearly,

$$\mathsf{deg}(G) \leq \mathsf{mad}(G).$$

With the new notions, repeating the proof of Theorem 1.4, we obtain that:

Lemma 1.6. *For a graph G,*

$$\chi(G) \leq \mathsf{deg}(G) + 1.$$

Three remarks are in place. First, due to Lemma 1.6, the degeneracy of G is sometimes called the *coloring number* of G. Second, there are numerous examples where the inequality in Lemma 1.6 is not tight: consider for example the biclique $K_{n,n}$ that is 2-colorable (i.e., bipartite) but has degeneracy n.

Third, the structure obtained in the proof of Lemma 1.6 is interesting on its own. If one looks algorithmically at a proof of Lemma 1.6 that is a copied almost verbatim proof of Theorem 1.4, an algorithm to $(\mathsf{deg}(G)+1)$-color a graph G is to first recursively "peel off" vertices of degree at most $\mathsf{deg}(G)$, ending with an empty graph, and while going back from the recursion color the peeled off vertex with a color not used by any of its already colored neighbors. This leads to the notion of a *degeneracy ordering* as the order in which we "peel off" vertices in the above process.

In this course we will be often interested in various properties of orderings of $V(G)$, so we introduce some notation. Formally, an ordering σ of $V(G)$ is a bijection from $V(G)$ to $[|V(G)|]$. We say that $v \in V(G)$ is *earlier*, *smaller*, or *to the left* of $u \in V(G)$ if $\sigma(v) < \sigma(u)$ and we denote it by $v <_\sigma u$. If $v <_\sigma u$, then we also say that u is *later*, *larger*, or *to the right* of v.

The set of *left neighbors* of $v \in V(G)$ in the ordering G is

$$\mathrm{Reach}_1[G, \sigma, v] = \{u \in V(G) \mid uv \in E(G) \wedge u <_\sigma v\}.$$

The *degeneracy of an ordering σ of $V(G)$* equals

$$\mathsf{deg}(G, \sigma) = \max_{v \in V(G)} |\mathrm{Reach}_1[G, \sigma, v]|.$$

Let us make a few observations about the new notions.

First, if one takes an order σ of $V(G)$ and iteratively for $i = 1, 2, \ldots, |V(G)|$ color the vertex $v := \sigma^{-1}(i)$ with a minimum color not appearing in $\mathrm{Reach}_1[G, \sigma, v]$, then one obtain a $(\mathsf{deg}(G, \sigma) + 1)$-coloring of G.

Second, for a given order σ of $V(G)$ and a subgraph $H \subseteq G$, if v is the σ-maximum element of $V(H)$, then $\mathrm{d}_H(v) \leq |\mathrm{Reach}_1[G, \sigma, v]|$. Hence, $\delta(H) \leq \mathsf{deg}(G, \sigma)$. Since the choice of H is arbitrary, $\mathsf{deg}(G) \leq \mathsf{deg}(G, \sigma)$.

Third, consider the following recursive procedure to compute an order σ of G: take a vertex $v \in V(G)$ of minimum degree, compute recursively an ordering σ' of $G - v$, and append v at the end of this order to obtain σ. Since at every step we pick a vertex of degree at most $\mathsf{deg}(G)$, this procedure results in an order σ with $\mathsf{deg}(G, \sigma) \leq \mathsf{deg}(G)$.

The last two observations prove the following.

Lemma 1.7.

$$\mathsf{deg}(G) = \min_\sigma \mathsf{deg}(G, \sigma).$$

Furthermore, an order attaining the minimum on the right hand side can be computed in linear time.

An order σ with $\mathsf{deg}(G, \sigma) = \mathsf{deg}(G)$ is sometimes called a *degeneracy ordering* of G.

With Lemma 1.7 in hand, we can now fully relate $\mathsf{deg}(G)$ and $\mathsf{mad}(G)$.

Lemma 1.8. *For every graph G,*

$$\mathsf{deg}(G) \leq \mathsf{mad}(G) \leq 2 \cdot \mathsf{deg}(G).$$

Proof. The first inequality was already discussed and follows directly from the fact that $\delta(H) \leq \mathsf{avgdeg}(H)$ for every $H \subseteq G$. For the second inequality, consider a degeneracy ordering σ of G. Then, for every $H \subseteq G$, the order σ induces an order σ_H of $V(H)$ with $\mathsf{deg}(H, \sigma_H) \leq \mathsf{deg}(G, \sigma)$ as for every $v \in V(H)$ we have $\mathrm{Reach}_1[H, \sigma_H, v] \subseteq \mathrm{Reach}_1[G, \sigma, v]$. However, then H has at most $|V(H)| \cdot \mathsf{deg}(H, \sigma_H)$ edges. The right inequality follows. □

Hence, while minimum and average degree of a graph can differ significantly, degeneracy and maximum average degree are closely related. In most of the reasonings in this course, the constant 2 in the statement of Lemma 1.8 does not matter at all, and thus we will most of the time use the notion of degeneracy and ditch maximum average degree.

1.3 Beyond degeneracy in planar graphs

In the language of degeneracy, Corollary 1.3 implies that planar graphs are 5-degenerate. Lemma 1.6 implies that they are 6-colorable. However, a celebrated 4-color theorem asserts that planar graphs are even 4-colorable, while an example of a 6-vertex clique shows that no all 5-degenerate graphs are even 5-colorable. Let us now discuss briefly what makes the class of planar graphs much more restricted than the class of 5-degenerate graphs.

First, the topology of the plane allows the well-known proof of the 5-colorability of planar graphs via the so-called *Kempe chains*. This argument — being about reachability in planar graphs within unbounded radius — is exactly the type of argument that we do *not* want to discuss at this course. In particular, we do not recall the Kempe chains argument here.

Second, note that an n-vertex clique with every edge subdivided once (i.e., a vertex is inserted in the middle of every edge of a clique) is 2-degenerate, but one cannot imagine any topological property that such a subdivided clique may admit.

Thus, *bounded degeneracy* is usually not a sufficient assumption to ensure nice algorithmic or structural properties of a graph class we consider. The theory of *sparsity* is about the following phenomenon: if one consider bounded degeneracy assumption at a slightly larger scale — at every constant radius (instead of the regular degeneracy being a degeneracy at radius one) — then one obtain a number of strong structural and algorithmic properties, while including a large number of graphs within the scope of the theory *and* retaining the simplicity of degeneracy-based arguments (as opposed to often nasty and full of traps and special cases topological arguments).

To give a flavor of a "slightly larger scale" sparsity phenomena, let us make one more exercise from discharging. (Two more exercises will be discussed at the tutorial.)

Lemma 1.9. *Every connected plane graph with minimum degree at least* 3 *contains a face and an incident vertex such that the sum of the legth of the face and the degree of the vertex is at most* 8.

Proof. We use discharging.
Setup. Every vertex and every face receives a charge of $+12$, every edge receives a charge of -12. The Euler's formula asserts that the total charge in the graph is $+24$.
First discharging phase. First, every vertex sends a charge of $+3$ to every incident edge and every face sends a charge of $+3$ to every incident edge. In this manner, every edge attains exactly zero charge.
Second discharging phase. Second, every vertex of degree 3 sends a charge of $+1$ to every incident face and every face of length 3 sends a charge of $+1$ to every incident vertex.

Final charge analysis. As discussed, every edge ends up with a zero charge. Consider now a vertex or a face v of degree/length k. By assumption, $k \geq 3$.

Case $k = 3$. Then v started with a charge of $+12$, lost $3 \cdot 3$ in the first phase and $3 \cdot 1$ in the second phase. Thus, v has nonpositive charge unless it has an incident face/vertex of length/degree 3 (that sent to v some charge in the second phase), which would prove the lemma.

Case $k \in \{4, 5\}$. Then v started with a charge of $+12$ and lost $4 \cdot 3$ or $5 \cdot 3$ in the first phase. Thus, again v has nonpositive charge unless it has an incident face/vertex of length/degree 3, which would prove the lemma.

Case $k \geq 6$. Then v started with a charge of $+12$, lost $k \cdot 3$ in the first phase, and gained at most k in the second phase. Hence, v ended up with nonpositive charge even if all its incident faces/vertices sent it a charge in the second phase.

Thus, a vertex or a face can have positive charge only if it is contained in the configuration whose existence is asserted by the lemma. Since the total charge is positive, the lemma is proven. □

1.4 Graphs excluding a (topological) minor

In this course we will be interested in properties of graph classes that are in some sense between planar graphs and graphs of bounded degeneracy. Let us now scout a number of classic examples of such graph classes. First, we need to understand what its mean to embed one graph into another topologically. This understanding is provided by the notion of a *minor*.

Definition 1.10. A graph H is a minor of G, written $H \preccurlyeq G$, if there is a *minor model* ϕ of H in G: a map ϕ which assigns to every vertex $v \in V(H)$ a connected subgraph $\phi(v) \subseteq G$ of G and to every edge $e \in E(H)$ an edge $\phi(e) \in E(G)$ such that

1. if $u, v \in V(H)$ with $u \neq v$ then $V(\phi(v)) \cap V(\phi(u)) = \emptyset$ and
2. if $e = uv \in E(H)$ then $\phi(e) = u'v' \in E(G)$ for vertices $u' \in V(\phi(u))$ and $v' \in V(\phi(v))$.

The set $\phi(v)$ for a vertex $v \in V(H)$ is called the *branch set* of v.

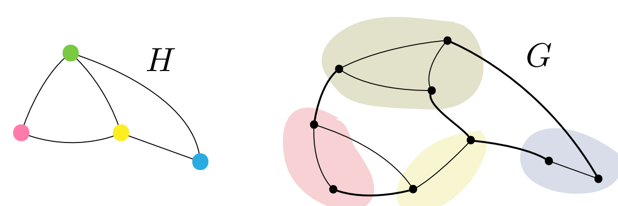

Figure 1: The graph H depicted above is a minor of the graph G. The branch sets $\phi(u) \subseteq V(G)$ of vertices $u \in V(H)$ are identified by colors in the figure.

Let G be a graph and uv be an edge of G. Define the *contraction* of uv as the operation of replacing both u and v by a single, new vertex w that is adjacent to all vertices that were previously adjacent to either u or v. Note that contraction does not introduce multiple edges (if x was adjacent to both u and v, then we put only one edge between x and w), thus we stay in the setting of simple graphs. It is easy to prove the following operational characterization of the minor order.

Lemma 1.11. *A graph H is a minor of a graph G if and only if H can be obtained from G by repeated use of the following operations: vertex deletion, edge deletion, and edge contraction.*

From the above lemma it trivially follows that if $J \preccurlyeq H$ and $H \preccurlyeq G$, then $J \preccurlyeq G$; a proof using the original definition is almost equally easy (essentially, compose the minor model mappings). Thus, the relation of being a minor imposes a partial order on graphs.

Definition 1.12. A graph H is a topological minor of G, written $H \preccurlyeq^{\text{top}} G$, if there is a *topological minor model* of H in G: a map ψ which assigns to every vertex $v \in V(H)$ a vertex $\phi(v) \in V(G)$ of G and to every edge $e \in E(H)$ a path $\psi(e)$ in G such that

1. if $u, v \in V(H)$ with $u \neq v$ then $\psi(v) \neq \psi(u)$,
2. if $e = uv \in E(H)$ then $\psi(e)$ is a path with endpoints u and v and
3. if $e, e' \in E(H)$ with $e \neq e'$, then $\psi(e)$ and $\psi(e')$ are internally vertex disjoint.

For a vertex $v \in V(H)$, the vertex $\psi(v)$ is the *pin* of v.

Again, it is easy to prove an operational characterization of the topological minor order.

Lemma 1.13. *A graph H is a topological minor of a graph G if and only if H can be obtained from G by repeated use of the following operations: vertex deletion, edge deletion, and contraction of an edge with one endpoint of degree 2.*

Clearly, if H is a topological minor of G, it is also a minor of G, but the implication in the other direction is far from true. Note that in a minor model ψ of H, the degree of $\psi(v)$ in G needs to be at least as large as the degree of v in H to accommodate all paths P_e starting in v. Hence, while a graph of maximum degree 3 can have arbitrarily large clique minors, it has no K_5 topological minor due to absence of vertices of degree more than 3.

We will be interested classes of graphs not containing one particular graph H as a minor or as a topological minor. If no graph from a class of graphs $\mathcal{G}$ contains H as a (topological) minor, then we say that $\mathcal{G}$ is *H-(topological)-minor-free.* Since we will be mostly interested in sparsity bounds for H-(topological)-minor-free graphs in terms of the size of H, we will be mostly focusing on the case $H = K_t$ for some integer t; note that if $H \subseteq H'$ then and H-(topological)-minor-free graph is trivially also H'-(topological)-minor-free.

Clearly, graphs of maximum degree t do not contain K_{t+2} as a topological minor. One can prove that the class of K_t-topological-minor-free graphs is "not much more" than the class of K_t-minor-free graphs "glued" with the class of graphs of maximum degree at most $t-2$. However, the proof of this fact and even the precise definitions of "not much more" and "glued" are far beyond the scope of this course.

Note that if G is planar, then any minor of G is also planar: it is straightforward to obtain a plane embedding of a minor of G from the embedding of G. The classic results of Kuratowski and Wagner show (topological)-minor-minimal graphs that are not planar.

Theorem 1.14 (Kuratowski, Wagner)**.** *The following conditions are equivalent for a graph G:*

1. *G is planar.*

2. *G does not contain K_5 nor $K_{3,3}$ as a minor.*

3. *G does not contain K_5 nor $K_{3,3}$ as a topological minor.*

Planar graphs are 5-degenerate. Does something similar hold for H-minor-free graphs? Yes, otherwise we probably would not discuss H-minor-free graphs in this course.

Theorem 1.15. *There exists a universal constant c such that for every $t \geq 2$ every K_t-minor-free graph has degeneracy at most $c \cdot t \cdot \sqrt{\log t}$.*

The proof of Theorem 1.15 is involved and technical, and thus we will resort to a weaker bound with a simpler proof.

Lemma 1.16. *For every $t \geq 2$ and $n \geq 1$, if an n-vertex graph does not contain K_t as a minor, then it has less than $2^t \cdot n$ edges.*

Proof. We use induction over $n + t$. For the base cases $t = 2$ and $n < t$ the claim is trivial. Otherwise, let G be an n-vertex graph with no K_t as a minor.

Pick $v \in V(G)$ and consider $G' \coloneqq G[N(v)]$. Since G has no K_t-minor, G' has no K_{t-1}-minor and thus by induction hypothesis G' has less than $2^{t-1} \cdot |N(v)|$ edges. Consequently, there exists $u \in N(v)$ that has degree less than 2^t in G', that is, $|N(v) \cap N(u)| < 2^t$. Then the contraction of the edge uv leads to a graph G'' with (i) $n-1$ vertices, (ii) at least $|E(G)| - 2^t$ edges, and (iii) no K_t-minor. Thus, the claim follows from the induction hypothesis applied to G''. □

1.5 Treewidth and treedepth

There are two important examples of minor-closed graph classes (i.e., graph classes closed under taking minors), related to *width parameters* of graphs. A *width parameter* is a function that assigns to every graph a nonnegative integer being the width of this graph. In this course, we will be using *treewidth* and *treedepth*.

Treedepth of a graph G, denoted $\mathtt{td}(G)$, can be defined recursively as follows.

1. Treedepth of an empty graph is 0.

2. Treedepth of a disconnected graph is the maximum of treedepth over the connected components of the graph.

3. Treedepth of a connected graph G equals $1 + \min_{v \in V(G)} \mathtt{td}(G - v)$.

In other words, treedepth is the minimum height of a recursive "elimination tree" of a graph, where at every step of the recursion we can delete one vertex and independently recurse on the connected components of the resulting graph. Equivalently, one can define $\mathtt{td}(G)$ as the minimum height of a rooted forest F such that there exists a bijection $\pi : V(G) \to V(F)$ such that for every $uv \in E(G)$, $\pi(u)$ and $\pi(v)$ are in ancestor-descendant relation in F. Such a pair (F, π) is called a *treedepth decomposition* of G and its width is the height of F (measured as the number of vertices on the longest root-to-leaf path).

Similarly as the degeneracy, treedepth can be defined with regards to the orderings of G. For a graph G, ordering σ of $V(G)$, and a vertex $v \in V(G)$, we define the *weakly reachable set of v*, denoted $\mathrm{WReach}_\infty[G,\sigma,v]$, as the set of those $u \in V(G)$ such that there exists a path P in G from u to v where u is the σ-minimum vertex of $V(P)$. Note that $v \in \mathrm{WReach}_\infty[G,\sigma,v]$. The *weak coloring number* of an order σ, denoted $\mathrm{wcol}_\infty(G,\sigma)$, is the maximum size of a weakly reachable set in this order, and the *weak coloring number* of a graph G is the minimum weak coloring number of an order of $V(G)$.

It turns out that the weak coloring number is *exactly* the treedepth of the graph.

Theorem 1.17. *For every graph G, it holds that*

$$\mathsf{td}(G) = \mathrm{wcol}_\infty(G).$$

Proof. We proceed by induction over the size of G. The statement is trivial for $|V(G)| \leq 1$.

If G is disconnected with connected components $G_1, G_2, \ldots, G_k$, by the induction hypothesis there exists an ordering σ_i of G_i with $\mathsf{td}(G_i) = \mathrm{wcol}_\infty(G_i) = \mathrm{wcol}_\infty(G_i,\sigma_i)$. By appending the orders σ_i one after another, we obtain an order σ of G with

$$\mathrm{wcol}_\infty(G,\sigma) = \max_{i=1}^{k} \mathrm{wcol}_\infty(G_i,\sigma_i) \leq \max_{i=1}^{k} \mathsf{td}(G_i) = \mathsf{td}(G).$$

Hence, $\mathrm{wcol}_\infty(G) \leq \mathsf{td}(G)$. In the other direction, every order σ' of $V(G)$ indudces and order σ'_i of $V(G_i)$ and $\mathrm{wcol}_\infty(G,\sigma') = \max_{i=1}^{k} \mathrm{wcol}_\infty(G_i,\sigma'_i)$. By the induction hypothesis, $\mathrm{wcol}_\infty(G_i,\sigma'_i) \geq \mathsf{td}(G_i)$ so $\mathrm{wcol}_\infty(G,\sigma') \geq \max_{i=1}^{k} \mathsf{td}(G_i) = \mathsf{td}(G)$, as desired.

If G is connected, let $v \in V(G)$ be such that $\mathsf{td}(G) = 1 + \mathsf{td}(G-v)$. By the induction hypothesis, let σ_1 be the order of $G-v$ with $\mathrm{wcol}_\infty(G-v,\sigma_1) = \mathrm{wcol}_\infty(G-v) = \mathsf{td}(G-v)$. Obtain an order σ of $V(G)$ by prepending v and the begining of the order σ_1. Then for every $u \in V(G) \setminus \{v\}$ we have $\mathrm{WReach}_\infty[G,\sigma,w] \subseteq \{v\} \cup \mathrm{WReach}_\infty[G-v,\sigma_1,w]$ and thus

$$\mathrm{wcol}_\infty(G) \leq 1 + \mathrm{wcol}_\infty(G-v,\sigma_1) \leq 1 + \mathsf{td}(G-v) = \mathsf{td}(G).$$

In the other direction, consider an order σ' of $V(G)$, and let w be the σ'-minimum vertex. By the connectivity of G, $w \in \mathrm{WReach}_\infty[G,\sigma',u]$ for every $u \in V(G)$. Thus, $\mathrm{wcol}_\infty(G,\sigma') = 1 + \mathrm{wcol}_\infty(G-w,\sigma'_1)$, where σ'_1 is the order σ' restricted to $G-w$. Hence,

$$\mathsf{td}(G) \leq 1 + \mathsf{td}(G-w) = 1 + \mathrm{wcol}_\infty(G-w) \leq 1 + \mathrm{wcol}_\infty(G-w,\sigma'_1) = \mathrm{wcol}_\infty(G,\sigma').$$

This finishes the proof of the lemma. ☐

We now move to a second width parameter, namely *treewidth.* A *tree decomposition* of a graph G consists of a tree T and a function $\beta : V(T) \to 2^{V(G)}$ such that:

1. For every $v \in V(G)$, the set $\{t \in V(T) \mid v \in \beta(t)\}$ is a nonempty connected subgraph of T.

2. For every $uv \in E(G)$ there exists $t \in V(T)$ with $u, v \in \beta(t)$.

The *width* of a tree decomposition (T,β) is $\max_{t \in V(T)} |\beta(t)| - 1$ and the *treewidth* of a graph G, denoted $\mathsf{tw}(G)$, is the minimum width of a tree decomposition of G.

It is straightforward that deleting a vertex or an edge cannot increase treewidth nor treedepth of the graph. It is simple (and postponed to the tutorials) to check that also contracting an edge cannot increase treewidth nor treedepth.

While being quite nonintuitive at the first glance, treewidth and tree decompositions play fundamental role in many branches of graph theory and graph algorithms. For example, the notion of treewidth properly formalizes an intuition that a planar graph of bounded radius cannot be too "thick".

Theorem 1.18. *A connected planar graph of radius r has treewidth at most $3r + 2$.*

Proof. Consider a plane embedding of G and let $v \in V(G)$ be such that every $u \in V(G)$ is within distance at most r from v. Since the treewidth cannot increase while adding an edge, we can assume that G is a triangulation, that is, every face is of length 3.

Let F be the set of faces of G. Let T be a shortest-path tree from v in G. For every edge $e \in E(G) \setminus E(T)$, its *dual* e^* is an edge connecting the two faces incident with e. Observe that $T^* := (F, \{e^* \mid e \in E(G) \setminus E(T))$ is a tree. For every $f \in F$, define $\beta(f)$ to be the union of the vertex sets of the three paths from u to v in T for all three vertices $u \in V(G)$ incident with f. Then a direct check shows that (T^*, β) is a tree decomposition of G of width at most $3r + 2$, as desired. □

Treewidth can again be casted into the language of vertex orderings. For an ordering σ of $V(G)$, we define the *elimination procedure* as follows: we iteratively take the σ-maximum vertex v of G, turn $N(v)$ into a clique, and delete v. The *width* of the elimination procedure is the maximum degree of a vertex upon its deletion, plus one.

The elimination procedure and its width can be also understood via the notions of *strongly reachable sets* and *strong coloring numbers*. For a graph G, ordering σ of $V(G)$, and a vertex $v \in V(G)$, the *strongly reachable set* of v, denoted $\mathrm{SReach}_\infty[G, \sigma, v]$, is defined as the set of those $u \in V(G)$ with $u \leq_\sigma v$ such that there exists a path P from v to u with all internal vertices to the right of v. Clearly, $v \in \mathrm{SReach}_\infty[G, \sigma, v] \subseteq \mathrm{WReach}_\infty[G, \sigma, v]$. The *strong coloring number* of an ordering σ, denoted $\mathrm{scol}_\infty(G, \sigma)$ is the maximum size of a strongly reachable set and the *strong coloring number* of a graph, denoted $\mathrm{scol}_\infty(G)$, is the minimum strong coloring number of an ordering of the vertex set.

Observe that given an order σ, the set $\mathrm{SReach}_\infty[G, \sigma, v]$ consists of v and all vertices u that are adjacent to v upon the deletion of v in the elimination procedure for σ. Thus, $\mathrm{scol}_\infty(G)$ is the minimum possible width of an elimination procedure for G.

Similarly as with treedepth and weak coloring number, the strong coloring number is the same as treewidth, barring the annoying "-1" in the definition of treewidth (that is there only to say that forests are of treewidth at most 1).

Theorem 1.19. *For every graph G, it holds that*

$$\mathrm{scol}_\infty(G) = \mathsf{tw}(G) + 1.$$

Proof. We use induction over the size of G. For $|V(G)| \leq 1$ the statement is trivial.

Consider first an order σ of $V(G)$ such that $\mathrm{scol}_\infty(G) = \mathrm{scol}_\infty(G, \sigma)$. Let v be the σ-maximum vertex. Let G' be the graph $G - v$ with $N(v)$ turned into clique and let σ' be the order σ' restricted to G'. By the choice of σ, $\mathrm{scol}_\infty(G', \sigma') \leq \mathrm{scol}_\infty(G, \sigma)$. By the induction hypothesis, $\mathsf{tw}(G') + 1 = \mathrm{scol}_\infty(G') \leq \mathrm{scol}_\infty(G', \sigma')$. That is, there exists a tree decomposition (T', β') of G' with $|\beta'(t)| \leq \mathrm{scol}_\infty(G, \sigma)$ for every $t \in V(T')$. Since $N(v)$ is a clique in G', there exists $t_0 \in V(T')$ with $N(v) \subseteq \beta'(t_0)$. Construct a tree decomposition (T, β) of G from (T', β') by appending a degree-1 neighbor t_1 to t_0 and setting $\beta(t_1) = \{v\} \cup N(v)$ and $\beta(t) = \beta'(t)$ for every $t \in V(T')$. This proves $\mathsf{tw}(G) + 1 \leq \mathrm{scol}_\infty(G)$.

In the other direction, let (T, β) be a tree decomposition of G of width $\mathsf{tw}(G) + 1$. Root T at an arbitrary node r and let $v \in V(G)$ be such that the distance from r to $\{t \in V(T) \mid v \in \beta(t)\}$ is maximum possible. Let G' be the graph obtained from G by turning $N(v)$ into a clique and deleting v.

Let $t_0 \in V(T)$ be the closest to r node of T with $v \in \beta(t_0)$. The choice of v and the properties of a tree decomposition imply that $N(v) \subseteq \beta(t_0)$. Consequently, if we define $\beta'(t) = \beta(t) \setminus \{v\}$ for every $t \in V(T)$, then (T, β') is a tree decomposition of G'. Hence, $\mathsf{tw}(G') \leq \mathsf{tw}(G)$.

By the induction hypothesis, there exists an ordering σ' of $V(G')$ with $\mathrm{scol}_\infty(G', \sigma') = \mathrm{scol}_\infty(G') = \mathsf{tw}(G') + 1$. Let σ be the order of $V(G)$ obtained from σ' by appending v on the right. Then, as $\{v\} \cup N(v) \subseteq \beta(t_0)$ and $\mathrm{SReach}_\infty[G, v, \sigma] = \{v\} \cup N(v)$, we have that $|\mathrm{SReach}_\infty[G, v, \sigma]| \leq |\beta(t_0)| \leq \mathsf{tw}(G) + 1$ while for every $u \in V(G')$ we have $\mathrm{SReach}_\infty[G', u, \sigma'] = \mathrm{SReach}_\infty[G, u, \sigma]$. This proves that $\mathrm{scol}_\infty(G) \leq \mathrm{scol}_\infty(G, \sigma) \leq tw(G) + 1$, as desired. □

The fact that a weakly reachable set is a superset of a strongly reachable set of a vertex in an ordering implies that $\mathrm{scol}_\infty(G) \leq \mathrm{wcol}_\infty(G)$ and thus

$$\mathsf{tw}(G) + 1 \leq \mathsf{td}(G).$$

As already discussed (and proven formally on the tutorial), if H is a minor of G, then $\mathsf{tw}(H) \leq \mathsf{tw}(G)$ and $\mathsf{td}(H) \leq \mathsf{td}(G)$. The notion of an elimination ordering and Theorem 1.19 immediately shows that

$$|E(G)| \leq |V(G)| \cdot \mathsf{tw}(G).$$

Hence, graphs of treewidth at most k are k-degenerate. Since $\mathsf{tw}(G)+1 \leq \mathsf{td}(G)$, graphs of treedepth at most k are $(k-1)$-degenerate.

We conclude with one remark regarding the notation. A reader at this point may wonder about the ∞ subscript in the notation $\mathrm{WReach}_\infty[G, \sigma, v]$, $\mathrm{SReach}_\infty[G, \sigma, v]$, wcol_∞, and scol_∞. This ∞ corresponds to no constraint on the length of the path P in the definition of weakly and strongly reachable sets. Later in the course, we will be studying *local* variants of weakly and strongly reachable sets with a bound on the length of the path P, and replace ∞ with this bound.

2 Bounded expansion and nowhere denseness

2.1 Definitions, examples, and main properties

The starting point of the theory of structural sparsity is the observation that sparsity understood as, say, boundedness of degeneracy is not really a robust notion when we would like to construct a theory of "locally sparse" structures. This is exemplified in the following observation.

Lemma 2.1. *For every graph G, the* 1-subdivision *of G — a graph obtained from G by replacing every edge by a 2-vertex path — is 2-degenerate.*

Proof. Let G' be the 1-subdivision of G. Then every subgraph of G' either contains a subdividing vertex — which has degree at most 2 — or consists only of original vertices of G, and hence is edgeless. □

From the point of view of theory-building, we are now making a project decision: do we consider a 1-subdivision of a large clique sparse or dense? If we decide that it is sparse, then we essentially end up with a theory based on degeneracy, arboricity, maximum average degree, etc., which is interesting, but not that deep. On the other hand, we can decide that a 1-subdivision of a large clique is dense, because a "dense structure" can be seen "at depth 1". Choosing this route leads to the theory of structural sparsity.

We first need to define containing a dense structure locally, at some bounded depth. For this, we will use the concept of *shallow minors*: minors suitably augmented with locality constraints.

Definition 2.2. The *radius* of a connected graph G is $\mathsf{rad}(G) = \min_{u \in V(G)} \max_{v \in V(G)} \mathrm{dist}(u, v)$.

Definition 2.3. Let H, G be graphs and let $d \in \mathbb{N}$. The graph H is a *depth-d minor of* G, written $H \preccurlyeq_d G$, if there is a minor model ϕ of H in G such that the branch set $\phi(v) \subseteq G$ has radius at most d for all $v \in V(G)$. We define

$$\nabla_d(G) \coloneqq \sup \left\{ \frac{|E(H)|}{|V(H)|} \ : \ H \preccurlyeq_d G \right\}$$

and

$$\omega_d(G) \coloneqq \sup\{t \ : \ K_t \preccurlyeq_d G\}.$$

Now that we defined the notion of local containment, we can think what is a "dense structure". When one thinks about it, there are two natural interpretations possible: either a graph of high edge density, or simply a clique. The first interpretation leads to the notion to graph classes of *bounded expansion*, and second to *nowhere dense* graph classes. These are the two central notions of the theory.

Definition 2.4. For a graph G and $d \in \mathbb{N}$ we define

$$\nabla_d(G) \coloneqq \sup \left\{ \frac{|E(H)|}{|V(H)|} \ : \ H \preccurlyeq_d G \right\} \qquad \text{and} \qquad \omega_d(G) \coloneqq \sup\{\, t \ : \ K_t \preccurlyeq_d G \,\}.$$

Further, for a class of graphs $\mathcal{C}$ and $d \in \mathbb{N}$, we define

$$\nabla_d(\mathcal{C}) \coloneqq \sup_{G \in \mathcal{C}} \nabla_d(G) \qquad \text{and} \qquad \omega_d(\mathcal{C}) \coloneqq \sup_{G \in \mathcal{C}} \omega_d(G).$$

In general, throughout this course, if π is a graph parameter, i.e., a function from graphs to nonnegative integers, then we may apply π to a graph class $\mathcal{C}$ by putting $\pi(\mathcal{C}) = \sup_{G \in \mathcal{C}} \pi(G)$. Then $\mathcal{C}$ has *bounded* π if $\pi(\mathcal{C})$ is finite.

Observe that the depth-0 minors of a graph are exactly its subgraphs, hence for every graph G we have $\nabla_0(G) = \frac{1}{2}\mathsf{mad}(G)$.

In bounded expansion we require all fixed-depth minors to have bounded average degree, as explained formally below.

Definition 2.5. A class $\mathcal{C}$ of graphs has *bounded expansion* if

$$\nabla_d(\mathcal{C}) < +\infty \qquad \text{for every } d \in \mathbb{N}.$$

Equivalently, there is a function $f \colon \mathbb{N} \to \mathbb{N}$ such that for all $d \in \mathbb{N}$ and $G \in \mathcal{C}$, we have $\nabla_d(G) \leq f(d)$.

In nowhere denseness we require fixed-depth minors to simply not contain all cliques.

Definition 2.6. A class $\mathcal{C}$ of graphs is *nowhere dense* if

$$\omega_d(\mathcal{C}) < +\infty \qquad \text{for every } d \in \mathbb{N}.$$

Equivalently, there is a function $t \colon \mathbb{N} \to \mathbb{N}$ such that for all $d \in \mathbb{N}$ and $G \in \mathcal{C}$, we have $\omega_d(G) \leq t(d)$.

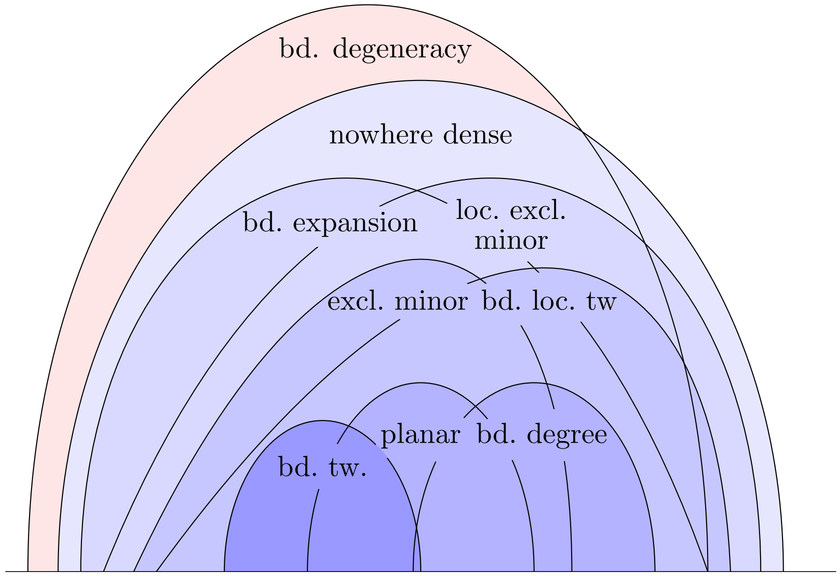

Figure 2: Illustration of the inclusions of sparse graph classes.

Note that every graph is a subgraph of some clique, so nowhere denseness is equivalent to requiring that for every depth $d \in \mathbb{N}$ we exclude at least one graph as a depth-d minor. A class that is not nowhere dense is called *somewhere dense*.

Clearly, every bounded expansion class is nowhere dense. Also, the class of all 2-degenerate graphs is neither nowhere dense nor it has bounded expansion, as witnessed by the example of a subdivided clique. It turns out, however, that many studied classes of sparse graphs, like planar graphs, classes with bounded genus or with bounded treewidth, classes with bounded maximum degree, or classes excluding a fixed minor, all have bounded expansion. See Figure 2.1 for an illustration of the inclusions of all the mentioned classes (and some more that we did not discuss). Let us now make this statement formal by proving that every class of bounded maximum degree and every class excluding a fixed minor has bounded expansion.

Lemma 2.7. *Let G be a graph of maximum degree Δ. Then $\nabla_d(G) \leq \Delta^{d+1}/2$. In particular, every graph class with bounded maximum degree has bounded expansion.*

Proof. We prove that every d-shallow minor of G has maximum degree at most Δ^{d+1}. Take a depth-d shallow minor model ϕ of some H in G, take any $u \in V(H)$, and fix some vertex $\gamma(u)$ of $\phi(u)$ such that all vertices of $\phi(u)$ are at distance at most d from $\gamma(u)$ in $\phi(u)$. Without loss of generality $\phi(u)$ is a tree of depth at most d, rooted at $\gamma(u)$. Construct $\overline{\phi(u)}$ from $\phi(u)$ by adding all edges $\phi(uv)$ for v ranging over neighbors of u in H; then $\overline{\phi(u)}$ is a tree of depth at most $d+1$ rooted at $\gamma(u)$. Observe that the number of neighbors of u in H is bounded by the number of leaves of $\overline{\phi(u)}$, which in turn is at most Δ^{d+1} because $\overline{\phi(u)}$ is a tree of depth at most $d+1$ with maximum degree at most Δ. □

Lemma 2.8. *Every class $\mathcal{C}$ that excludes some fixed graph H as a minor (i.e. no graph from $\mathcal{C}$ contains H as a minor) has bounded expansion.*

Proof. If H has t vertices, then $\mathcal{C}$ in particular excluded K_t as a minor. By Lemma 1.16, we infer that $\frac{|E(G)|}{|V(G)|} \leq 2^t$ for every $G \in \mathcal{C}$. Furthermore, every minor of a K_t-minor-free graph is K_t-minor-free, so this bounds holds also for all G ranging over minors of graphs from $\mathcal{C}$, independent of depth. We conclude that $\nabla_d(\mathcal{C}) \leq 2^t$ for every $d \in \mathbb{N}$. □

Finally, there is a question whether there are classes that are nowhere dense, but have unbounded expansion. This is indeed true, but no really natural examples are known. To give an example of such a class, we introduce one more concept.

Definition 2.9. The *girth* of a graph G, denoted $\mathsf{girth}(G)$, is the length of a shortest cycle in G.

It is not hard to construct large graphs of high girth.

Lemma 2.10. *For every positive integer k there exists a simple graph G that has maximum degree at most k and girth at least k, but contains at least $\frac{k|V(G)|}{4}$ edges.*

Proof. Without loss of generality $k \geq 3$. We construct the graph G as follows. Start with a set $V(G)$ consisting of $4 \cdot k^{k-1}$ vertices and with no edges. Then add edges to the graph as long as possible, but not violating any of the following conditions: the girth of G is at least k and the maximum degree of G is at most k. It suffices to prove that at the end of this procedure, G has at least $k|V(G)|/4$ edges. Take any vertex v of G. By a similar reasoning as in the proof of Lemma 2.7, the number of vertices at distance at most $k-1$ from v is at most $1 + k + k^2 + \ldots + k^{k-1} \leq 2 \cdot k^{k-1}$. As $|V(G)| = 4 \cdot k^{k-1}$, there are at least $2 \cdot k^{k-1} = |V(G)|/2$ vertices of G that are at distance at least k from v. Each such vertex w must have degree k in G, for otherwise we could add edge vw to G. Hence G contains at least $|V(G)|/2$ vertices of degree k, implying that $|E(G)| \geq k|V(G)|/4$. □

The following lemma provides essentially the only known example of a nowhere dense class without bounded expansion.

Lemma 2.11. *The class $\mathcal{C} = \{G : \Delta(G) \leq \mathsf{girth}(G)\}$ is nowhere dense and has unbounded expansion.*

Proof. We first prove that the class is nowhere dense. Assume that $K_t \preccurlyeq_d G$ for some d, $t \geq 3$, and $G \in \mathcal{C}$. Fix a depth-d minor model of K_t in G. Then this model contains a cycle of length at most $6d + 3$, implying that $\mathsf{girth}(G) \leq 6d + 3$. By the definition of $\mathcal{C}$ we conclude that the maximum degree of G is at most $6d + 3$. As in the proof of Lemma 2.7, every depth-d minor of G has maximum degree at most $(6d + 3)^{d+1}$, which implies that $t \leq 1 + (6d + 3)^{d+1}$. We conclude that for every $d \in \mathbb{N}$, we have $\omega_d(\mathcal{C}) \leq \max(3, 1 + (6d + 3)^{d+1})$, so $\mathcal{C}$ is indeed nowhere dense.

Second, to see that $\mathcal{C}$ has unbounded expansion, note that Lemma 2.10 asserts that $\mathcal{C}$ contains graphs of unbounded average degree. □

Reducts of classes. As we observed before, a minor of a minor is a minor. We now prove that also a shallow minor of a shallow minor is a shallow minor, where the parameters are adjusted accordingly.

Lemma 2.12. *Suppose $J \preccurlyeq_b H$ and $H \preccurlyeq_a G$ for some values of $a, b \in \mathbb{N}$. Then $J \preccurlyeq_{2ab+a+b} G$.*

Proof. Fix models ϕ_J of J in H and ϕ_H of H in G witnessing $J \preccurlyeq_b H$ and $H \preccurlyeq_a G$. We define a minor model ϕ of J in G as follows. For every $u \in V(J)$, define $\phi(u)$ to be the union of subgraphs $\phi_H(v)$ for $v \in \phi_J(u)$ and edges $\phi_H(vv')$ for $vv' \in \phi_J(u)$. Further, for every $uu' \in E(J)$ define $\phi(uu') = \phi_H(\phi_J(uu'))$. It is straightforward to see that ϕ defined in this manner is a depth-$(2ab + a + b)$ minor model of J in G. Indeed, if for $u \in V(J)$ we first take $v \in V(\phi_J(u))$ certifying that $\mathsf{rad}(\phi_J(u)) \leq b$, and then we take $w \in V(\phi_H(v))$ witnessing that $\mathsf{rad}(\phi_H(v)) \leq a$, then w is at distance at most $2ab + a + b$ from every vertex of $\phi(u)$ within this branch set; see Figure 3 for an illustration. □

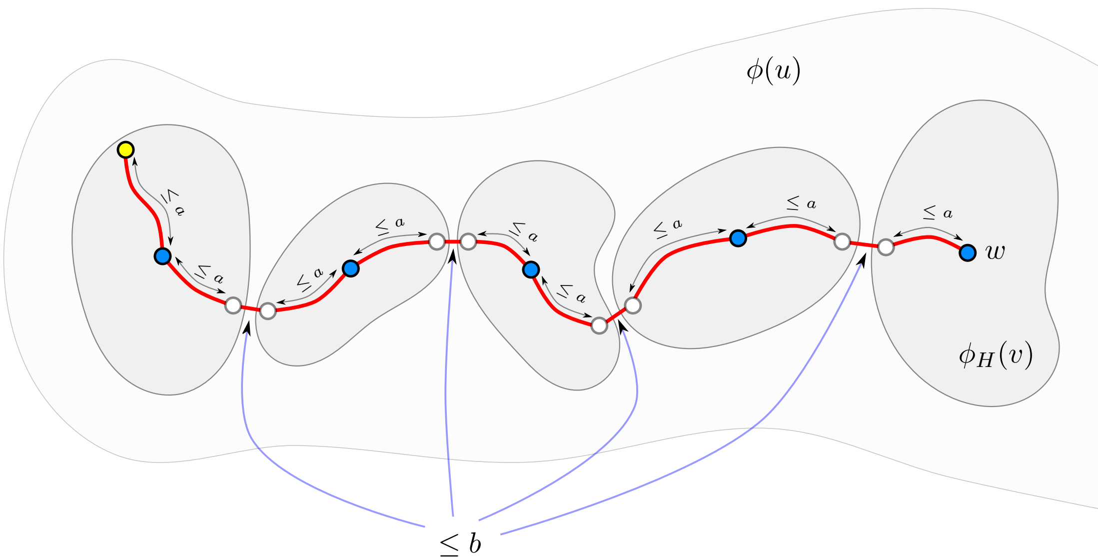

Figure 3: Proof that $\mathsf{rad}(\phi(u)) \leq 2ab + a + b$ for each $u \in V(J)$. Blue vertices witness that corresponding branch sets in ϕ_H have radius at most a. The red path witnesses that an arbitrary vertex of $\phi(u)$ (depicted in yellow) is at distance at most $2ab + a + b$ from w within this graph.

The above observation motivates looking at taking shallow minors as kind of an operator on classes of graphs.

Definition 2.13. For a graph class $\mathcal{C}$ and $d \in \mathbb{N}$, we define the *depth-d reduct* of $\mathcal{C}$ as:

$$\mathcal{C} \triangledown d = \{H \ : \ H \text{ is a depth-}d \text{ minor of some } G \in \mathcal{C}\}.$$

Corollary 2.14. *If $\mathcal{C}$ is a graph class and $a, b \in \mathbb{N}$, then*

$$(\mathcal{C} \triangledown a) \triangledown b \subseteq \mathcal{C} \triangledown (2ab + a + b).$$

Consequently,

$$\nabla_b(\mathcal{C} \triangledown a) \leq \nabla_{2ab+a+b}(\mathcal{C}) \qquad \text{and} \qquad \omega_b(\mathcal{C} \triangledown a) \leq \omega_{2ab+a+b}(\mathcal{C}).$$

In particular, if $\mathcal{C}$ has bounded expansion (resp. is nowhere dense), then for every fixed $b \in \mathbb{N}$, the class $\mathcal{C} \triangledown b$ also has bounded expansion (resp. is nowhere dense).

2.2 Topological minors

In the definitions introduced the previous section we relied on the minor order as the underlying notion of topological embeddings for graphs. However, recall that in the previous section we also defined a quite different notion of embedding: *topological minor*. We could easily define the whole theory using shallow topological minors, instead of minors.

Definition 2.15. A graph H is a *topological depth-d minor of* G, written $H \preccurlyeq_d^{\mathrm{top}} G$, if there is a model ϕ of H in G such that the paths $\phi(e)$ have length at most $2d+1$ for all $e \in E(H)$.

In this definition we use length $2d+1$ to reflect the basic relation with the minor order: $H \preccurlyeq_d^{\mathrm{top}} G$ entails $H \preccurlyeq_d G$.

Definition 2.16. Fix $d \in \mathbb{N}$. For a graph G we define

$$\widetilde{\nabla}_d(G) \coloneqq \sup\left\{\frac{|E(H)|}{|V(H)|} \ :\ H \preccurlyeq_d^{\mathrm{top}} G\right\} \qquad \text{and} \qquad \widetilde{\omega}_d(G) \coloneqq \sup\left\{t\ :\ K_t \preccurlyeq_d^{\mathrm{top}} G\right\}.$$

and for a graph class $\mathcal{C}$ we define

$$\widetilde{\nabla}_d(\mathcal{C}) \coloneqq \sup_{G\in\mathcal{C}} \widetilde{\nabla}_d(G) \qquad \text{and} \qquad \widetilde{\omega}_d(\mathcal{C}) \coloneqq \sup_{G\in\mathcal{C}} \widetilde{\omega}_d(G).$$

A class of graphs has *bounded topological expansion* if $\widetilde{\nabla}_d(\mathcal{C}) \leq +\infty$ for all $d \in \mathbb{N}$, and is *topologically nowhere dense* if $\widetilde{\omega}_d(\mathcal{C}) \leq +\infty$ for all $d \in \mathbb{N}$.

Recall that without the assumption of bounded depth, the minor and topological minor orders are quite different. It turns out, however, that by relying on topological minors instead of minors, we define *exactly* the same notions.

Theorem 2.17. *Every graph class is nowhere dense if and only if it is topologically nowhere dense.*

Theorem 2.18. *Every graph class has bounded expansion if and only if it has bounded topological expansion.*

2.2.1 Nowhere denseness via topological minors

Theorem 2.17 is a direct corollary of the following lemma. Indeed, the finiteness of $\{\omega_d(\mathcal{C})\}_{d\in\mathbb{N}}$ trivially implies the finiteness of $\{\widetilde{\omega}_d(\mathcal{C})\}_{d\in\mathbb{N}}$, while the reverse implication is provided below.

Lemma 2.19. *Suppose $\mathcal{C}$ is a graph class such that for some $d \in \mathbb{N}$, $\mathcal{C}$ admits all cliques as depth-d minors; that is, every clique is a depth-d minor of some graph from $\mathcal{C}$. Then $\mathcal{C}$ admits all cliques as topological depth-$(3d+1)$ minors.*

Proof. Fix any $t \in \mathbb{N}$. We would like to show that K_t is a topological depth-d minor of some graph from $\mathcal{C}$. Let us select a large constants k and s depending on d and t as follows:

$$k \coloneqq t^2 \qquad \text{and} \qquad s \coloneqq 2 + k^{d+1}.$$

The rationale behind this choice of s will become clear in the course of the proof; for now the reader should think of it as "something very large compared to t".

Since $\mathcal{C}$ admits every clique as a depth-d minor, there is some $G \in \mathcal{C}$ for which we can find a depth-d minor model ϕ of K_s in G. We shall prove that G also admits K_t as a topological depth-d minor. For this we will use parts of ϕ to construct an appropriate topological minor model.

The intuition behind the first step is that in a topological minor model of K_t, we necessarily need to have vertices of degree $t-1$. We find them in the branch sets of ϕ using the fact that each of these branch sets is essentially a bounded-depth tree with very many leaves.

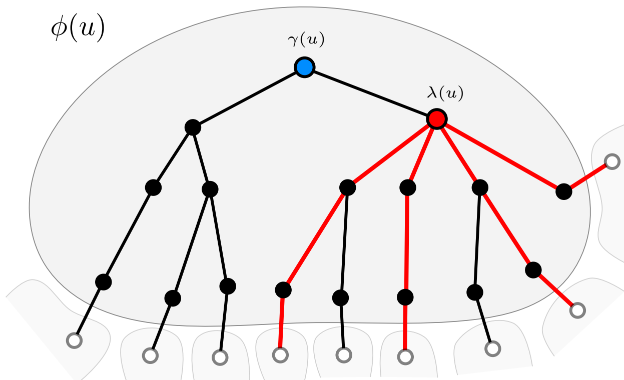

Figure 4: Choosing $\lambda(u)$ and $\mathcal{P}(u)$ (the latter depicted in red).

More precisely, similarly as in Lemma 2.7, for each $u \in V(K_s)$ we constuct $\overline{\phi(u)}$ from $\phi(u)$ by adding, for every $v \in V(K_s)$ with $u \neq v$, the edge $\phi(uv)$; denote the set of added edges by $F(u) \coloneqq \{\phi(uv) \colon v \in V(K_s),\, v \neq u\}$. Fix any vertex $\gamma(u) \in V(\phi(u))$ such that every vertex of $\phi(u)$ is at distance at most d from $\gamma(u)$. Then by restricting $\phi(u)$ to a minimal shortest-paths tree containing all the edges of $F(u)$, we can assume without loss of generality that $\overline{\phi(u)}$ is a tree of depth at most $d+1$ rooted at $\gamma(u)$ where the leaves of $\overline{\phi(u)}$ are precisely the endpoints of edges of $F(u)$ not residing in $\phi(u)$. Since the size of $F(u)$ is $s-1$, which is larger than k^{d+1}, and the depth of $\overline{\phi(u)}$ is at most $d+1$, it follows that in $\overline{\phi(u)}$ we can find a vertex $\lambda(u)$ that has more than k children. For each of the subtrees rooted in the children of $\lambda(u)$, fix one path contained in this subtree that starts in $\lambda(u)$ and ends in an edge of $F(u)$. Let $\mathcal{P}(u)$ be the family of those paths; note that they pairwise meet only at u. Thus, each path of $\mathcal{P}(u)$ connects $\lambda(u)$ to a vertex of $\phi(v)$ for some $v \neq u$, and traverses only edges of $\phi(u)$ plus the edge $\phi(uv)$. Every vertex $v \neq u$ of K_s for which $\phi(v)$ is connected in this way to $\lambda(u)$ by a path from $\mathcal{P}(u)$ will be called a *buddy* of u. Thus, every vertex $u \in V(K_s)$ has more than k buddies, but the relation of being a buddy is not necessarily symmetric.

Let us pick an arbitrary vertex subset $A \subseteq V(K_s)$ of size t; vertices $\lambda(u)$ for $u \in A$ will be the images of vertices of K_t in our topological depth-$(3d+1)$ minor model of K_t in G. We would like to assign to every $u \in A$ a set of its $t-1$ *private buddies* such that the following conditions are satisfied: every private buddy does not belong to A, and every vertex $v \notin A$ is a private buddy of at most one vertex $u \in A$. Since every vertex of A has more than $k = t^2$ buddies, this can be done greedily as follows. One by one, every vertex $u \in A$ picks $t-1$ of its buddies to be its private buddies, where the selection is done among buddies of u that do not belong to A and that were not selected by vertices of A considered earlier. Since $|A| = t$, the number of buddies excluded from selection in this manner is at most $(t-1)+(t-1)^2 = t(t-1)$, which leaves more than $k-t(t-1) = t$ buddies to choose from.

Now, for every vertex $u \in A$ we assign the $t-1$ private buddies of u to the $t-1$ vertices v of K_t other than u. Let $w_{u,uv}$ be the private buddy of u assigned to $v \neq u$. For each unordered pair $\{u,v\} \subseteq A$ with $u \neq v$, construct a path $P(uv)$ in G as follows (see Figure 5):

- Starting from $\lambda(u)$, follow the path belonging to $\mathcal{P}(u)$ that leads from $\lambda(u)$ to $\phi(uw_{u,uv})$.
- The graph $\phi(w_{u,uv})$ is connected and has radius at most d, hence there is a path of length at most $2d$ within $\phi(w_{u,uv})$ that connects the endpoints of $\phi(uw_{u,uv})$ and $\phi(w_{u,uv}w_{v,uv})$ that belong to $\phi(w_{u,uv})$. Follow this path to the latter endpoint.
- Traverse the edge $\phi(w_{u,uv}w_{v,uv})$.
- Similarly as above, there is a path of length at most dr within $\phi(w_{v,uv})$ that connects the endpoints of $\phi(w_{u,uv}w_{v,uv})$ and $\phi(w_{u,uv}v)$ that belong to $\phi(w_{v,uv})$. Follow this path to the latter endpoint.
- Finally, follow the path belonging to $\mathcal{P}(v)$ that leads from the edge $\phi(vw_{u,uv})$ to $\lambda(v)$.

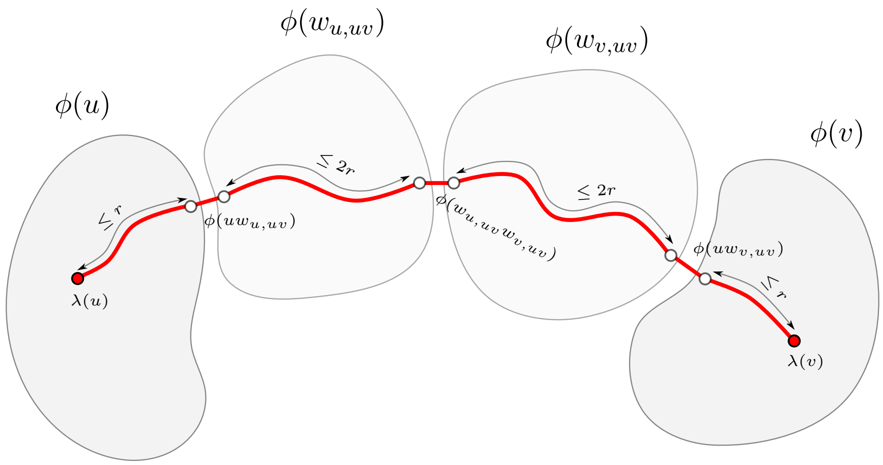

Figure 5: Construction of the path $P(uv)$ (in red).

Observe that $P(uv)$ is a path of length at most $6d+3$ connecting $\lambda(u)$ with $\lambda(v)$. Moreover, from the construction it readily follows that paths $P(uv)$ for different $\{u,v\} \subseteq A$ are internally vertex-disjoint. We conclude that mapping vertices of K_t to vertices $\lambda(u)$ for $u \in A$, and edges of K_t to appropriate paths $P(uv)$ for $\{u,v\} \subseteq A$, constitutes a topological depth-$(3d+1)$ minor model of K_t in G. $\square$

Let us note that the proof of Lemma 2.19 actually yields a slightly stronger conclusion that describes the dependencies between parameters.

Corollary 2.20. *For every graph G and $d \in \mathbb{N}$, it holds that*

$$\widetilde{\omega}_d(G) \leq \omega_d(G) \leq 1 + (\widetilde{\omega}_{3d+1}(G)+1)^{2d+2}.$$

It is possible to bound $\omega_d(G)$ also by a function of $\widetilde{\omega}_d(G)$ (i.e. with the same radius d), but this is technically more challenging.

2.2.2 Bounded expansion via topological minors

Again, Theorem 2.18 is a direct corollary of the following lemma.

Lemma 2.21. *For every graph G and $d \in \mathbb{N}$, we have*

$$\widetilde{\nabla}_d(G) \leq \nabla_d(G) \leq 2^{d^2+3d+3} \cdot (\lceil \widetilde{\nabla}_d(G) \rceil)^{(d+2)^2}.$$

The proof of Lemma 2.21 is beyond the scope of this course, due to its technicality, and was not included in the lectures. However, its conclusion, in particular Theorem 2.18, is important for the theory. For completeness and for interested readers, we include a proof of Lemma 2.21 below; it spans the remainder of this section.

We will need the following classic lemma stating that in any graph we can always find a bipartite subgraph that contains at least half of the edge set.

Lemma 2.22. *Let G be a graph with m edges. Then G contains a bipartite subgraph $H \subseteq G$ with $V(H) = V(G)$ and $|E(H)| \geq m/2$.*

Proof. Construct a partition (X, Y) of $V(G)$ as follows: each vertex u is placed in X with probability $\frac{1}{2}$ and in Y with probability $\frac{1}{2}$, and these random choices are taken independently. Let H be the subgraph of G obtained by preserving all vertices and only those edges whose endpoints belong to different sides of the partition (X, Y); clearly H is bipartite. Observe that the probability that a fixed edge uv is preserved in H is equal to $\frac{1}{2}$. By linearity of expectation, the expected number of edges in H is equal to $\frac{m}{2}$. Therefore, there exists at least one choice of the partition (X, Y), and thus at least one choice of the subgraph H, for which H has at least $m/2$ edges. $\square$

Let us now recall some intuition from the proof of Lemma 2.19. Suppose ϕ is a depth-d minor model ϕ of some graph H in some graph G. For a vertex $u \in V(H)$, let $\overline{\phi(u)}$ be constructed from $\phi(u)$ by adding all edges $\phi(uv)$ for neighbors v of u in H. The endpoints of these additional edges that are not contained in $\phi(u)$ will be called the *legs* of $\overline{\phi(u)}$. We will say that u is a *spider* (of depth at most d) in ϕ if $\overline{\phi(u)}$ is a star with every edge subdivided at most d times (i.e., replaced by a path of length at most $d + 1$), and the leaves of $\overline{\phi(u)}$ are exactly its legs. Observe that if all vertices of H are spiders of depth at most d in ϕ, then ϕ can be trivially turned into a depth-d topological minor model H in G.

Therefore, intuitively our goal in the proof of Lemma 2.21 is the following: starting with a depth-d minor model ϕ of some really dense graph H in G, transform it into a depth-d minor model ϕ' of some graph H' in G such that H' is still quite dense, but all the vertices of H' are spiders in ϕ'. Recall that in the proof of Lemma 2.19, within each branch set of the considered shallow minor model we essentially found one spider, and this was enough for our purposes. Here, the situation will be more complicated, as we will extract multiple spiders from one branch set; the number of extracted spiders will be linear in the degree of the corresponding vertex. The essence of this argument is encapsulated in the following lemma, see Figure 6 for an illustration.

Lemma 2.23. *Suppose T is a rooted tree of depth at most $d + 1$ and with ℓ leaves, and let a be a positive integer. Then one can find a subset Z of internal nodes of T and a partial function α from leaves of T to Z with the following properties:*

- *For each $z \in Z$, $\alpha^{-1}(z)$ is a set of more than a leaves that are all descendants of z.*

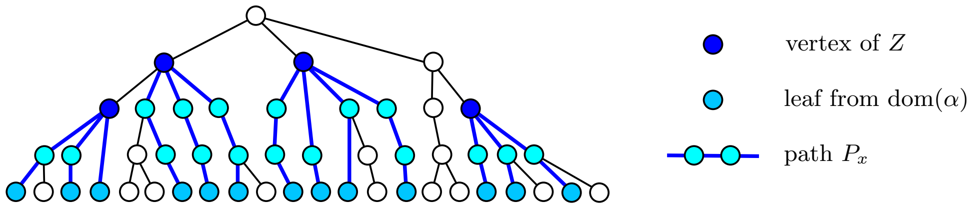

Figure 6: An example tree of depth 4 and a possible outcome of applying Lemma 2.23 on it, for $a = 2$. Each leaf $x \in \mathrm{dom}(\alpha)$ is mapped by α to the vertex of Z to which the path P_x leads.

- *If for a leaf $x \in \mathrm{dom}(\alpha)$ by P_x we denote the unique path in T from x to $\alpha(x)$, then for all distinct $x, y \in \mathrm{dom}(\alpha)$ we have $V(P_x) \cap V(P_y) = \{\alpha(x)\} \cap \{\alpha(y)\}$.*

- *The domain of α has size at least $\lfloor \frac{\ell-1}{a^{d+1}} \rfloor$.*

Proof. We apply induction on the number of leaves ℓ. For the base case, observe that for $\ell \le a^{d+1}$ we may take Z and α to be empty, thus from now on we may assume that $\ell > a^{d+1}$.

As T has depth $d+1$ and more than a^{d+1} leaves, there is at least one node in T that has more than a children. Let z be a deepest (i.e., furthest from the root) node in Z that has this property; say z has $k > a$ children. Then each subtree rooted at a child of z has depth at most $d+1$, and every node residing in it has at most a children. Therefore, each such subtree contains at most a^{d+1} leaves, implying that there are at most $k \cdot a^{d+1}$ leaves that are descendants of z in total.

Construct a tree T' from T by taking every leaf of T that is not a descendant of z, and including in T' the whole path from it to the root; i.e., T' is the union of these paths. Thus, the set of leaves of T' is equal to the set of leaves of T minus those leaves that are descendants of z, and the subtree of T rooted at z is completely removed in T'. Apply induction hypothesis to T', yielding a set Z' and partial function α' from the leaves of T' to Z'. Now put $Z = Z' \cup \{z\}$ and construct α by extending α' as follows: for each subtree rooted at a child of z, pick any leaf from this subtree and map it to z. Thus $|\alpha^{-1}(z)| = k > a$ and the leaves of $\alpha^{-1}(z)$ reside in pairwise different subtrees rooted at the children of z; by combining this with the induction assumption, it is straightforward to see that α satisfies the first two required conditions. To see that the domain of α has size at least $\lfloor \frac{\ell-1}{a^{d+1}} \rfloor$, one may use the induction assumption plus the observation that for each subtree rooted at a child of z, the domain of α contains one of the at most a^{d+1} leaves of this subtree. ◻

From Lemma 2.23 we deduce the following statement, where we replace many branch sets with spiders in one shot.

Lemma 2.24. *Suppose H is a bipartite graph, say with bipartition (X, Y), whose edge density is more than $(2a)^{d+2}$, where a is a positive integer. Let ϕ be a depth-d minor model of H in some graph G. Then there exists a bipartite graph H' with bipartition (X', Y), and a depth-d minor model ϕ' of H' in G, with the following properties:*

- *the edge density of H' is more than a;*

- *each $u \in X'$ is a spider of depth at most d in ϕ'; and*

- *for each $v \in Y$, we have $\phi'(v) = \phi(v)$.*

Proof. Similarly as in the proof of Lemma 2.19, we may assume that each branch set $\phi(u)$ is a tree of depth at most d rooted at some center vertex $\gamma(u)$, whose leaves are exactly those endpoints of edges $\phi(uv)$ for $uv \in E(H)$ that reside in $\phi(u)$. Indeed, within $\phi(u)$ there is a tree with these properties, and we may drop from $\phi(u)$ all the features (vertices and edges) outside this tree. Then each $\overline{\phi(u)}$ is a tree of depth at most $d+1$ rooted at $\gamma(u)$, and the leaves of $\overline{\phi(u)}$ are exactly its legs.

Now, for each vertex $u \in X$ we apply Lemma 2.23 to the tree $\overline{\phi(u)}$, for parameter $2a$. This yields a set $Z_u \subseteq V(\phi(u))$ and a partial map α_u from the legs of $\overline{\phi(u)}$ to Z_u. Let $X' = \sum_{u \in X} Z_u$. We construct a graph H' on the vertex set $X' \uplus Y$ as follows: for each $w \in X'$, say $w \in Z_u$ for some $u \in X$, let us inspect the set of legs $\alpha_u^{-1}(w)$ of $\overline{\phi(u)}$. For each leg $\ell \in \alpha_u^{-1}(w)$, take the vertex $v \in Y$ such that $\ell \in V(\phi(v))$, and make w adjacent to v in H'. We may construct a depth-d minor model ϕ' of H' in G as follows. First, for $v \in Y$ put $\phi'(v) = \phi(v)$ and for $w \in X'$, say $w \in Z_u$ for some $u \in X$, construct $\phi'(w)$ by taking w and adding, for each leg $\ell \in \alpha_u^{-1}(w)$, the path in $\overline{\phi(u)}$ between w and ℓ with the last edge (incident to the leg) removed. It is straightforward to see that the model ϕ' has all the required properties, in particular every vertex $w \in X'$ is a spider of depth at most d in ϕ'.

We are left with making sure that the edge density of H' is more than a. For a vertex $u \in X$, let $\mathrm{d}_H(u)$ be the degree of u in H; then $|E(H)| = \sum_{u \in X} \mathrm{d}_H(u)$. By Lemma 2.23, we have that each vertex of X' has degree more than $2a$ in H', thus

$$|E(H')| > 2a|X'|. \tag{1}$$

On the other hand, by Lemma 2.23 again, we have that each vertex $u \in X$ gives rise to at least $\lfloor \frac{\mathrm{d}_H(u)-1}{(2a)^{d+1}} \rfloor \geq \frac{\mathrm{d}_H(u)}{(2a)^{d+1}} - 1$ edges in H'. Therefore, we have

$$|E(H')| \geq \sum_{u \in X} \left(\frac{\mathrm{d}_H(u)}{(2a)^{d+1}} - 1 \right) = \frac{|E(H)|}{(2a)^{d+1}} - |X| \geq \frac{(2a)^{d+2}}{(2a)^{d+1}} (|X| + |Y|) - |X| \geq 2a|Y|. \tag{2}$$

By combining (1) and (2) we infer that

$$|X'| + |Y| < \left(\frac{1}{2a} + \frac{1}{2a} \right) \cdot |E(H')| = \frac{|E(H')|}{a},$$

which means that the edge density of H' is more than a. □

With all the tools prepared, we may now move to the main proof.

Proof of Lemma 2.21. Let $a := \lceil \widetilde{\nabla}_d(G) \rceil$, and let function f be defined as $f(x) = (2x)^{d+2}$; then we need to prove that $\nabla_d(G) \leq 2 \cdot f(f(a))$. For the sake of contradiction, suppose there is a depth-d minor H_0 of G with edge density more than $2 \cdot f(f(a))$. By Lemma 2.22, H_0 contains a bipartite subgraph H_1 with edge density more than $f(f(a))$; obviously H_1 is still a depth-d minor of G. Let (X, Y) be a bipartition of H_1, and let ϕ_1 be a depth-d minor model of H_1 in G.

We apply Lemma 2.24 to H_1 and its model ϕ_1 in G, with parameter $f(a)$. This yields a bipartite graph H_2 with bipartition (X', Y), and a depth-d minor ϕ_2 of H_2 in G, such that: the edge density of H_2 is more than $f(a)$, all vertices of X' are spiders in ϕ_2, and all vertices of Y have the same branch sets in ϕ_2 as in ϕ_1. Next, again apply Lemma 2.24, this time to H_2 and its model ϕ_2 in G, and with parameter a, but with the sides reversed: we treat side Y to be X in the lemma statement, and side X' to be Y in the lemma statement. This yields a bipartite graph H_3 with bipartition (X', Y'), and a depth-d minor ϕ_3 of H_3 in G, such that: the edge density of H_3 is more than a, all vertices of Y' are spiders in ϕ_3, and all vertices of X' have the same branch sets in ϕ_3 as in ϕ_2. Recall that all vertices of X' were spiders in ϕ_2. In ϕ_3, for each $v \in X'$ we still have that $\overline{\phi_3(v)}$ is a star with every edge replaced by a path of length at most $d+1$, but some of these paths may no longer lead to the legs of $\overline{\phi_3(v)}$ due to dropping some connections in the construction of H_3 and ϕ_3. We may, however, just remove these unnecessary paths in each $\phi_3(v)$ to make every vertex of X' a spider in ϕ_3.

After this modification, every vertex of $X' \cup Y' = V(H_3)$ is a spider of depth at most d in ϕ_3, so as we argued before, the model ϕ_3 in fact witnesses that H_3 is a depth-d topological minor of G. However, the edge density of H_3 is larger than $a = \lceil \widetilde{\nabla}_d(G) \rceil$, a contradiction. □

2.3 Shallow congested minors

So far we have seen that bounded expansion means sparsity of shallow minors and shallow topological minors. We now see that we can also think about minors with *congestion*: instead of requiring the branch sets to be disjoint, we allow them to overlap, but we bound the number of branch sets that can overlap at one vertex.

Definition 2.25. We say that a graph H is a *depth-d minor with ply c* of a graph G if there exists a mapping ϕ from vertices of H to connected subgraphs of G such that:

- for every $uv \in E(H)$, the subgraph $\phi(u)$ and $\phi(v)$ *touch*: they either share a vertex or there is an edge in G with one endpoint in $\phi(u)$ and second in $\phi(v)$;
- for every $u \in V(H)$, the subgraph $\phi(u)$ has radius at most d; and
- for every $w \in V(G)$, there are at most c vertices $u \in V(H)$ satisfying $w \in \phi(u)$.

Obviously, depth-d minors of ply 1 are exactly standard depth-d minors. Also, the same reasoning as in Lemma 2.12 yields the following.

Lemma 2.26. *Suppose J is a depth-a minor with ply k of H and H is a depth b minor with ply ℓ of G. Then J is a depth-$(2ab + a + b)$ minor with ply $k\ell$ of G.*

We now show that bounded-depth bounded-ply minors of a sparse graph are also sparse.

Lemma 2.27. *Let G be a graph and let H be a depth-t ply-c minor of G. Then for each $d \in \mathbb{N}$,*

$$\nabla_d(H) \leq c + 2c^2(2dt + d + t + 1)^2 \cdot \nabla_{2dt+d+t}(G).$$

Proof. By Lemma 2.26, every depth-d minor of H is actually a depth-$(2dt + d + t)$ minor with ply c of G. Hence, it suffices to prove the lemma for $d = 0$, that is, prove that the edge density in H is bounded as follows:

$$\frac{|E(H)|}{|V(H)|} \leq c + 2c^2(t+1)^2 \cdot \nabla_t(G).$$

Indeed, to estimate the density of depth-d minors of H we will just use the above inequality but with parameter t set to $2dt + d + t$.

Let ϕ be a depth-d ply-c minor model of H in G. For a vertex $v \in V(G)$, by $\phi^{-1}(v)$ we denote the set of those $h \in V(H)$, for which $v \in \phi(h)$. For each $h \in V(H)$, fix any vertex $\gamma(h) \in \phi(h)$ that is at distance at most t from each vertex of $\phi(h)$.

Call an edge $hh' \in E(H)$ *degenerate* if either $\gamma(h) \in \phi(h')$ or $\gamma(h') \in \phi(h)$. Observe that each vertex $v \in V(G)$ may give rise to at most $|\gamma^{-1}(v)| \cdot |\phi^{-1}(v)|$ degenerate edges in H, while $|\phi^{-1}(v)| \leq c$ for each $v \in V(G)$. Hence, the total number of degenerate edges is at most

$$\sum_{v \in V(G)} |\gamma^{-1}(v)| \cdot |\phi^{-1}(v)| \leq c \cdot \sum_{v \in V(G)} |\gamma^{-1}(v)| = c|V(H)|.$$

For each nondegenerate edge $hh' \in E(H)$, fix any path $P_{hh'}$ of length at most $2t+1$ connecting $\gamma(h)$ and $\gamma(h')$ such that $P_{hh'}$ can be decomposed into a prefix of length at most t that is entirely contained in $\phi(h)$ and a suffix of length at most t that is entirely contained in $\phi(h')$.

Draw a linear order $<$ on $V(H)$ uniformly at random. We shall say that a vertex $h \in V(H)$ is *dominant* if h is $<$-largest among $\phi^{-1}(\gamma(h))$. Further, we shall say that an edge $hh' \in E(H)$ is *dominant* if it is non-degenerate and within the set $\bigcup_{w \in P_{hh'}} \phi^{-1}(w)$, vertices h and h' are the two $<$-largest elements. Observe that if an edge is dominant, then both its endpoints are dominant as well. Let J be a subgraph of H consisting of all dominant vertices and all dominant edges. We now prove that in expectation J is a dense subgraph of H, however J is also a depth-t minor of G (with ply 1).

Claim 1. *We have*

$$\mathbb{E}\left[|E(J)|\right] \geq (|E(H)| - c|V(H)|) \cdot \frac{1}{2c^2(t+1)^2}.$$

Proof. Since the total number of degenerate edges is at most $c|V(H)|$, by linearity of expectation it suffices to prove that every non-degenerate edge is dominant with probability at least $\frac{1}{2c^2(t+1)^2}$. For this, observe that since ϕ has ply at most c, we have

$$\left| \bigcup_{w \in P_{hh'}} \phi^{-1}(w) \right| \leq 2(t+1)(c-1)+2.$$

Then the probability that h and h' are the two $<$-largest elements of this set is at least

$$\frac{1}{\binom{2(t+1)(c-1)+2}{2}} \geq \frac{1}{2c^2(t+1)^2},$$

as claimed. ⌟

Claim 2. *J is a depth-t minor of G.*

Proof. For $h \in V(H)$, let $X(h)$ be the set of those vertices u of G, for which h is the $<$-largest element of $\phi^{-1}(u)$. Observe that if h is dominant, then $\gamma(h) \in X(h)$. Define a mapping ψ from vertices of J to connected subgraphs of G as follows: for a dominant vertex h, we let $\psi(h)$ be the subgraph of $G[X(h)]$ induced by all vertices at distance at most t from $\gamma(h)$ in $G[X(h)]$. Observe that subgraphs $\{\psi(h)\}_{h \in V(J)}$ are pairwise disjoint and each of them has radius at most t. Furthermore, if hh' is a dominant edge, then $P_{hh'}$ can be decomposed into a prefix of length at most t that is entirely contained in $X(h)$ and a suffix of length at most t that is entirely contained in $X(h')$. Then these prefix and suffix are respectively contained in $\psi(h)$ and $\psi(h')$, hence the edge on $P_{hh'}$ between them is an edge connecting a vertex of $\psi(h)$ with a vertex of $\psi(h')$. We conclude that ψ is a depth-t minor model of J in G. ⌟

By Claim 1, there exists an ordering $<$ of $V(H)$ for which J satisfies

$$|E(J)| \geq (|E(H)| - c|V(H)|) \cdot \frac{1}{2c^2(t+1)^2}.$$

On the other hand, by Claim 2 we have

$$\nabla_t(G) \geq \frac{|E(J)|}{|V(J)|} \geq \frac{|E(J)|}{|V(H)|}.$$

By combining these two inequalities we conclude that

$$\frac{|E(H)|}{|V(H)|} \leq c + 2c^2(t+1)^2 \cdot \frac{|E(J)|}{|V(H)|} \leq c + 2c^2(t+1)^2 \cdot \nabla_t(G),$$

as required. □

For a class of graphs $\mathcal{C}$ and $c, t \in \mathbb{N}$, we define

$$\mathcal{C} \nabla^c t = \{H : H \text{ is a depth-}t \text{ minor with ply at most } c \text{ of a graph from } \mathcal{C}\}.$$

Further, we shall say that a class $\mathcal{C}$ has *polynomial expansion* of degree k if $\nabla_d(\mathcal{C}) \leq p(d)$ for a polynomial $p(\cdot)$ of degree k. Then Lemma 2.27 directly implies the following.

Corollary 2.28. *Fix $c, t \in \mathbb{N}$. Then for every class of bounded expansion $\mathcal{C}$, the class $\mathcal{C} \nabla^c t$ also has bounded expansion. Futher, if $\mathcal{C}$ has polynomial expansion of degree k, then $\mathcal{C} \nabla^c t$ has polynomial expansion of degree $k+2$.*

3 Edge density in nowhere dense classes

In classes of bounded expansion we have a linear number of edges even after applying contractions of subgraphs of any fixed constant radius. A natural question is whether a similar characterization can be given for nowhere dense classes, which are defined qualitatively (by exclusion of cliques as shallow minors) rather than quantitatively. It turns out that the answer is affirmative. More precisely, we will prove the following theorem.

Theorem 3.1. *Suppose $\mathcal{C}$ is a nowhere dense class of graphs. Then for every $r \in \mathbb{N}$ and every $\varepsilon > 0$ there exists a constant $N = N(r, \varepsilon)$ such that for every graph $G \in \mathcal{C} \nabla r$ with $n \geq N$ vertices, we have that G has less than $n^{1+\varepsilon}$ edges.*

Observe that Theorem 3.1 is equivalent to saying that there is a function $f(r, \varepsilon)$ such that every graph $G \in \mathcal{C} \nabla r$ has at most $f(r, \varepsilon)|V(G)|^{1+\varepsilon}$ edges, without any lower bound on its size. Intuitively, the number of edges may be super-linear in the number of vertices, but it is as close to linear as we would like (this is sometimes called *almost linear*). Also, observe that Theorem 3.1 provides a dichotomy, for if $\mathcal{C}$ is somewhere dense, then for some r it contains all cliques as depth-r minors, where the number of edges grows quadratically. Thus, there is no "middle ground": every graph class $\mathcal{C}$ either at some depth r admits graphs in which the number of edges grows quadratically in the number of vertices, or on all levels this growth is almost linear.

The proof of Theorem 3.1 that we are going to present is due to Zdenek Dvořák. It gives worse bounds than the best known, but it is relatively easy. The main idea is to give a "densification procedure": given an n-vertex graph with at least $n^{1+\varepsilon}$ edges, we will find either a large depth-1 clique as its depth-1 minor (contradicting the fact that the graph is drawn from a nowhere dense class), or a suitably large depth-1 minor that has n' vertices and $(n')^{1+\varepsilon+\varepsilon^2}$ edges. By iterating this $\mathcal{O}(1/\varepsilon^2)$ times (so a constant number of times) we will eventually reach a contradiction, as the number of edges of a graph would exceed the trivial quadratic upper bound.

We first need to prepare a simple auxiliary lemma.

Lemma 3.2. *Suppose G is an n-vertex graph with at least dn edges, for some $d \in \mathbb{N}$. Then G contains a subgraph G' with minimum degree at least d and $|V(G')| \geq \sqrt{n}$.*

Proof. Take the graph G and as long as it contains a vertex of degree smaller than d, delete this vertex. Obviously, the graph cannot eventually become empty in this way, because then it would have at most $(d-1)n$ edges. Let G' be the obtained graph; by construction, G' has minimum degree at least d. Further, observe that while removing vertices of $V(G) - V(G')$, we deleted at most $(d-1)n$ edges of the graph. If we had $|V(G')| < \sqrt{n}$, then the number of remaining edges would be at most $\binom{\sqrt{n}}{2}$, which is smaller than n, so in total in the graph we would have less than dn edges; a contradiction. □

Also, we will use the following classic variant of Chernoff's bound.

Theorem 3.3 (Chernoff's bound)**.** *Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli random variables with success probability p. Let $S = X_1 + \ldots + X_n$ and let $\mu = np$ be the expected value of S. Then for every $\delta \in (0,1)$, we have*

$$\mathsf{Pr}(|S - \mu| > \delta\mu) \leq 2\exp\left(-\frac{\delta^2\mu}{3}\right).$$

We are ready to give the main "engine" of the proof, namely the densification lemma.

Lemma 3.4. *For every $\varepsilon \in (0, \frac{2}{3}]$ we can find $M(\varepsilon) \in \mathbb{N}$, with $M(\cdot)$ being a non-increasing function, such that the following holds. Suppose G is a graph on $n \geq M(\varepsilon)$ vertices with minimum degree at least n^ε. Then either G contains K_t as a depth-1 minor for some $t \geq \log n$, or we can find a depth-1 minor G' of G with $n' \geq n^{1-\varepsilon}$ vertices and at least $(n')^{1+\varepsilon+\varepsilon^2}$ edges.*

Proof. Select a vertex subset $A \subseteq V(G)$ by picking every vertex of G independently with probability $\frac{2\log n}{n^\varepsilon}$ (by taking $M(\varepsilon)$ to be large enough, we may assume that this is smaller than 1). The expected size of A is $2n^{1-\varepsilon}\log n$, so by Chernoff's bound, we have

$$\mathsf{Pr}(n^{1-\varepsilon}\log n \leq |A| \leq 3n^{1-\varepsilon}\log n) \geq 1 - 2\exp\left(-\frac{n^{1-\varepsilon}\log n}{6}\right). \tag{3}$$

Observe that we may take $M(\varepsilon)$ large enough so that for $n > M(\varepsilon)$ this probability is at least $\frac{2}{3}$.

Let B be the set of vertices outside of A that have at least $\log n$ neighbors in A. We now estimate the size of B. Take any vertex u, by the assumption about minimum degree we have that u has at least n^ε neighbors in G. Each of these neighbors is selected to A independently with probability $\frac{2\log n}{n^\varepsilon}$, hence the expected number of neighbors of u in A is at least $2\log n$. By Chernoff's bound again, we have that the probability that u has less than $\log n$ neighbors in A is bounded as follows:

$$\mathsf{Pr}(|N(u) \cap A| < \log n) \leq 2\exp\left(-\frac{\log n}{6}\right) = 2n^{-\frac{1}{6}}.$$

In addition, independently of this, u can be selected to A with probability $\frac{2\log n}{n^\varepsilon}$. This means that u is excluded from B with probability

$$\Pr(u \notin B) \leq 2n^{-\frac{1}{6}} + \frac{2\log n}{n^\varepsilon},$$

which is at most $\frac{1}{6}$ for large enough n (i.e., we put $M(\varepsilon)$ large enough so that for $n > M(\varepsilon)$ this holds). Therefore, the expected size of $V(G) - B$ is at most $n/6$, so by Markov's inequality we have

$$\Pr\left(|B| \geq \frac{n}{2}\right) \geq \frac{2}{3}. \tag{4}$$

Now, perform the following greedy process. Start with defining G' to be the bipartite graph induced between A and B in G. Then, iterate through vertices of B one by one and inspect their neighborhoods. Let b be the next vertex of B to be considered. Inspect the neighborhood of b in A (in the graph G'); this neighborhood is of size at least $\log n$. If it is not a clique, take any nonedge aa' in this neighborhood and contract b onto a; thus we add the edge aa' to $G'[A]$, and possibly more edges. If it is a clique, then $G'[A]$ contains a clique of size $\log n$. It can be easily seen that the previous operations were only contracting vertices of B onto vertices of A, so in total G' stays a depth-1 minor of G. Hence, we have found K_t for some $t \geq \log n$ as a depth-1 minor of G, which can be reported as the outcome of the lemma.

We are left with considering the situation when the process succeeds in contracting all vertices of B onto A. Thus, G' is a 1-shallow minor of G and its vertex set is A. In (3) and (4) we argued that, provided $M(\varepsilon)$ is chosen large enough, each of the following events happens with probability at least $\frac{2}{3}$:

$$n^{1-\varepsilon}\log n \leq |A| \leq 3n^{1-\varepsilon}\log n \qquad \text{and} \qquad |B| \geq \frac{n}{2}.$$

Hence, both of them happen with probability at least $\frac{1}{3}$. From now on we assume that this is the case, as for at least one experiment such A and B can be found.

Observe that every contraction of a vertex of B onto A introduced one new edge to G', hence G' has at least $|B|$ edges and exactly $|A|$ vertices. Therefore, the edge density in G' can be lower bounded as follows:

$$\frac{|E(G')|}{|V(G')|} \geq \frac{|B|}{|A|} \geq \frac{\frac{n}{2}}{3n^{1-\varepsilon}\log n} = \frac{n^\varepsilon}{6\log n}. \tag{5}$$

However, G' has $n' = |A| \leq 3n^{1-\varepsilon}\log n$ vertices, so

$$n \geq \left(\frac{n'}{3\log n}\right)^{\frac{1}{1-\varepsilon}}. \tag{6}$$

Putting (5) and (6) together, and using the assumption that $\varepsilon \leq \frac{2}{3}$, we infer that

$$\frac{|E(G')|}{|V(G')|} \geq \frac{(n')^{\frac{\varepsilon}{1-\varepsilon}}}{54\log^3 n}$$

Now observe that

$$\frac{\varepsilon}{1-\varepsilon} = \varepsilon(1 + \varepsilon + \varepsilon^2 + \ldots) \geq \varepsilon + \varepsilon^2 + \varepsilon^3.$$

Hence, we have

$$\frac{|E(G')|}{|V(G')|} \geq \frac{(n')^{\varepsilon+\varepsilon^2+\varepsilon^3}}{54\log^3 n}.$$

Recall that we have $n' \geq n^{1-\varepsilon}\log n$. Hence, by taking $M(\varepsilon)$ large enough we can guarantee that $(n')^{\varepsilon^3} \geq 54\log^3 n$. Therefore, we conclude that

$$\frac{|E(G')|}{|V(G')|} \geq (n')^{\varepsilon+\varepsilon^2}.$$

Together with $|V(G')| = n' \geq n^{1-\varepsilon}$, this concludes the proof. □

With all the tools prepared, we can wrap up the proof of Theorem 3.1.

Proof of Theorem 3.1. Observe that since the class $\mathcal{C}\nabla r$ is also nowhere dense (Corollary 2.14), it suffices to prove the statement for $r = 0$. Further, without loss of generality we may assume that $\mathcal{C}$ is closed under taking subraphs. This means that it suffices to prove that for each fixed $\varepsilon > 0$, there exists $N = N(\varepsilon)$ such that every graph $G \in \mathcal{C}$ with $n > N$ vertices has less than $n^{1+\varepsilon}$ edges.

Fix $\varepsilon > 0$ and suppose that $G \in \mathcal{C}$ has n vertices and at least $n^{1+\varepsilon}$ edges. We will appply a reasoning to G that leads to a contradiction, but in order for it to applicable, we need that the number of vertices of G is large enough. This "large enough" yields the sought lower bound $N(\varepsilon)$.

Starting with $G_0 \coloneqq G$ and $\varepsilon_0 \coloneqq \varepsilon$, we construct a sequence of graphs $G_0, G_1, G_2, \ldots$ and a sequence of reals $\varepsilon_0 < \varepsilon_1 < \varepsilon_2 < \ldots$, all smaller than $\frac{2}{3}$, as follows. Suppose G_{i-1} is already defined, and let n_{i-1} be the number of its vertices; we maintain the invariant that G_{i-1} has at least $n_{i-1}^{1+\varepsilon_{i-1}}$ edges. Apply Lemma 3.2 to G_{i-1}, yielding its subgraph G'_{i-1} that has at least $\sqrt{n_{i-1}}$ vertices and minimum degree at least $n_{i-1}^{\varepsilon_{i-1}}$. Next, apply Lemma 3.4 to G'_{i-1}, assuming for a moment that G'_{i-1} has enough vertices for this to be possible. If this application yields K_t as a depth-1 minor of G'_{i-1} for some $t \geq \log n_{i-1}$, we stop the construction; as we will see in a moment, assuming the initial vertex count n is large enough, this yields a contradiction. Otherwise, we obtain a graph G_i which is a depth-1 minor of G'_{i-1} and has the following properties: G_i has $n_i \geq (n_{i-1})^{\frac{1-\varepsilon}{2}} \geq (n_{i-1})^{\frac{1}{6}}$ vertices and at least $(n_i)^{1+\varepsilon_i}$ edges, where $\varepsilon_i \coloneqq \varepsilon_{i-1} + \varepsilon_{i-1}^2$. Finally, if it turns out that $\varepsilon_i > \frac{2}{3}$, we stop the construction and apply the same procedure as above one more time to G_i, but using parameter $\frac{2}{3}$ instead of ε_i. Observe that this application has to yield K_t as a depth-1 minor of G'_i for some $t \geq \log n_i$, because the other outcome would be a graph with a super-quadratic number of edges.

The above procedure, if applicable, terminates after at most $k \coloneqq \lceil \frac{1}{\varepsilon^2} \rceil$ iterations, for in each iteration we have that ε_i is larger by at least $\varepsilon_{i-1}^2 \geq \varepsilon^2$ than ε_{i-1}. Further, G_1 is a depth-1 minor of G_0, G_2 is a depth-1 minor of G_1, and so on, so by a straightforward induction using Lemma 2.12 we get that G_i is a depth-d_i minor of G for $d_i \coloneqq \frac{3^i-1}{2}$. Let t be such that K_t is not a depth-d_{k+1} minor of any graph from $\mathcal{C}$; such t exists by the assumption that $\mathcal{C}$ is nowhere dense.

We now need to make sure that provided n is large enough, in the above procedure the following two conditions are satisfied: first, at each iteration G_i has size at least $M(\varepsilon_i)$, so that Lemma 3.4 is applicable, and second, if this application yields a large clique as a depth-1 minor, then it is indeed a contradiction with $G \in \mathcal{C}$. For the first condition, note that $M(\varepsilon_i) \leq M(\varepsilon)$ for all i, and in each iteration the number of vertices of the next graph is at least the sixth root of the number of vertices of the previous one. Hence, it suffices to assume that $n > M(\varepsilon)^{6^k}$ to ensure that Lemma 3.4 is always applicable throughout the procedure. For the second condition, observe that we eventually obtain a clique of size at least $\log n^{\frac{1}{6^{k+1}}}$ as a minor at depth at most d_{k+1}. Hence, it suffices to assume that $n > 2^{t \cdot 6^{k+1}}$ to make sure that this contradicts the assumption that $G \in \mathcal{C}$. The maximum of the two above bounds may be set as $N(\varepsilon)$. $\square$