Chapter 1: Measuring sparsity

Compilation date: October 24, 2017

1 Introduction

1.1 Overview

The theory of bounded expansion and nowhere dense graph classes is a young but rapidly maturing subject. Nowhere denseness provides a very robust notion of uniform sparseness in graphs and many familiar families of sparse graph classes are nowhere dense. These include for example graphs of bounded tree-width, planar graphs, graphs of bounded genus and graphs that exclude a minor or topological minor.

The development of the theory of bounded expansion and nowhere dense graph classes is strongly driven by algorithmic questions. This line of research is based on the observation that many problems, such as the dominating set problem, which are considered intractable in general, can be solved efficiently on restricted graph classes, e.g. on the above mentioned classes of graphs of bounded tree-width, planar graphs, and graph classes excluding a fixed (topological) minor. Many algorithmic results that were first established for specific graph classes can be generalized to nowhere dense classes, and most interestingly, it turns out that nowhere dense classes form a natural limit for many algorithmic techniques for sparse graph classes.

In this course we are going to present the structural and algorithmic theory of nowhere dense graph classes.

1.2 Sparse graphs

Bounded average degree. A first idea to capture the concept of sparse classes of graphs is to consider graphs of bounded average degree. Let G be an undirected, simple graph. For a vertex $v \in V(G)$ we write $d_G(v)$ (or simply $d(v)$) for the degree of v in G, that is, for its number of neighbors.

Definition 1. The *average degree* of a graph G is the number

$$\mathsf{avg}(G) = \frac{\sum_{v \in V(G)} d(v)}{|V(G)|} = \frac{2|E(G)|}{|V(G)|}.$$

A class $\mathcal{C}$ of graphs is said to have bounded average degree if there is a constant d such that $\mathsf{avg}(G) \leqslant d$ for all $G \in \mathcal{C}$.

From an algorithmic perspective, this definition of sparsity is not sufficiently robust. Observe that given any graph G of order $n := |V(G)|$, we can bound its average degree by *padding*, that is, simply by adding n^2 isolated vertices. While this reduces the average degree below 2, for many problems it does not significantly change the structure of the graph. For instance, adding extra isolated vertices does change the problem of finding a minimum vertex cover. Therefore, to prevent padding arguments of this form, we may want to require that sparse graphs remain sparse if we take a subgraph, i.e. that sparse graphs do not have dense subgraphs.

Degeneracy. Hence we would like to say that a graph has *uniformly bounded average degree d* if all of its subgraphs have average degree at most d. This definition is up to constant factors captured by the following well-studied concept of degeneracy.

Definition 2. A graph G is *d-degenerate* if every subgraph of G contains a vertex of degree at most d. The *degeneracy* of G, denoted $\mathsf{deg}(G)$, is the minimum d such that G is d-degenerate.

Degeneracy can be also seen structurally via the following notion of *degeneracy ordering.*

Definition 3. Suppose $\sigma = (v_1, \dots, v_n)$ is an ordering of the vertices of a graph G. The *degeneracy* of σ, denoted $\mathsf{deg}(G, \sigma)$, is the maximum over vertices v_i of the number of neighbors of v_i that are smaller in σ. That is: $\mathsf{deg}(G, \sigma) = \max_{i=1,\dots,n} |\{j : j < i \text{ and } v_j v_i \in E(G)\}|$.

Lemma 1. *A graph is d-degenerate if and only if it has a vertex ordering of degeneracy at most d.*

Proof. Let G be the graph in question. From right to left, suppose G admits a vertex ordering σ of degeneracy at most d. Let H be any subgraph of G. Then the largest vertex of H with respect to σ has degree at most d in H, as it is adjacent to at most d vertices that are smaller in σ.

From left to right, we prove by induction on $n := |V(G)|$ that such an ordering exists. If $n \leqslant d$, there is nothing to show. Now assume that G has $n > d$ vertices and assume that the claim holds for all d-degenerate graphs with $n-1$ vertices. As G is d-degenerate, in G there exists a vertex v of degree at most d. Note that every subgraph of a d-degenerate graph is again d-degenerate. Hence, by assumption, we can order the vertices of $G - v$ as $v_1, \dots, v_{n-1}$ with the desired properties. Defining $v_n := v$ gives us an ordering of $V(G)$ with the desired properties. □

Observe that if an n-vertex graph admits a vertex ordering of degeneracy at most d ($d \geqslant 1$), then it has less than $d(n-1)$ edges. This is because we can assign each edge to the larger (in the ordering) endpoint, and thus each vertex, apart from the first one in the ordering, is assigned at most d edges.

Corollary 2. *An n-vertex d-degenerate graph has at most $d(n-1)$ edges.*

For $d \geqslant 1$ we have $d(n-1) < dn$, and this latter bound will be commonly used: a d-degenerate graph has less than dn edges, provided $d \geqslant 1$. We can now show that degeneracy captures exactly what we wanted to define with the notion of uniformly bounded average degree.

Lemma 3. *Let G be a graph. Then*

$$\mathsf{deg}(G) \leqslant \max_{H \subseteq G} \mathsf{avg}(H) \leqslant 2\mathsf{deg}(G).$$

Proof. For the first inequality, observe that a graph of average degree at most d must contain a vertex of degree at most d; apply this observation to every subgraph of G. For the second inequality, take any subgraph H of G, and observe that by Corollary 2 we have that $\mathsf{avg}(H) = 2|E(H)|/|V(H)| \leqslant 2\mathsf{deg}(H) \leqslant 2\mathsf{deg}(G)$. □

We now briefly discuss some algorithmic aspects of degeneracy. The argument contained in the proof of Lemma 1 can be easily turned into a quadratic algorithm that computes a vertex ordering of a graph of optimum degeneracy: just extract minimum-degree vertices one by one, thus constructing the ordering from right to left. With some effort (tutorials), one can improve the running time to linear.

Lemma 4. *There exists an algorithm that given a graph G with n vertices and m edges, in time $\mathcal{O}(n+m)$ computes its vertex ordering of degeneracy equal to $\mathsf{deg}(G)$.*

We will usually assume that graphs are given to our algorithm via adjacency lists. Observe that having computed an optimum degeneracy ordering σ, in time $\mathcal{O}(n+m)$, we can further scan the adjacency list of each vertex u and construct the sublist consisting of the at most d neighbors that are smaller than u in σ. Now, to check whether two vertices u and v are adjacent it suffices to check if u appears on this sublist of v or vice versa; this is because one of them is larger than the other in the degeneracy ordering. Thus, after linear-time preprocessing, the adjacency checks in a d-degenerate graph can be performed in time $\mathcal{O}(d)$.

Let us consider a simple example how to use this uniform sparsity condition algorithmically.

Theorem 5. *Let $d \geqslant 1$ be an integer. There is an algorithm which, given a d-degenerate n-vertex graph G and an integer $k \geqslant 1$ as input, decides whether G contains a clique of size at least k in time $\mathcal{O}(k^2 d^k \cdot n)$.*

Proof. In a pre-processing step we compute a vertex ordering $\sigma = (v_1, \ldots, v_n)$ of G of degeneracy at most d. This takes time $\mathcal{O}(n+m) = \mathcal{O}(dn)$.

Now we guess the largest vertex v of the k-clique we are looking for (that is, we iterate through all vertices), and test whether among $k-1$ of its at most d smaller neighbours we can find the remaining clique vertices. For this, we simply enumerate all $(k-1)$-tuples from its list of smaller neighbours (time $\mathcal{O}(d^{k-1})$) and test whether they form a clique (time $\mathcal{O}(k^2 \cdot d)$). □

We say that a parameterized decision problem is *fixed-parameter tractable* with respect to a parameter k if there exists a computable function f and a constant c such that on an input instance of length n we can decide in time $f(k) \cdot n^c$ whether the instance belongs to the problem. Hence, we have just proved that the clique problem parameterized by the size of the solution is fixed-parameter tractable on graphs of bounded degeneracy.

Hence for the clique problem, the concept of degeneracy is very useful, however, for other problems it is not universally satisfactory. For example, we can make every graph 2-degenerate by subdividing every edge once. Here, by subdividing an edge we mean the operation of replacing an edge $\{u, v\}$ by a path of length 2. For various problems such as the *distance*-2 *dominating set problem*, subdividing edges does not change the nature of the graph significantly. Hence, in addition to closure under subgraphs, we may want to require of our notion of sparsity that it should be invariant under such subdivisions or local modifications. These two requirements together exactly yield the concept of *bounded expansion* and *nowhere denseness.*

2 Bounded expansion and nowhere dense classes of graphs

Definitions and examples. The following definition of a *minor* is a key concept in graph structure theory. Intuitively, it formalizes the idea of embedding one graph into the other topologically.

Definition 4. A graph H is a minor of G, written $H \preccurlyeq G$, if there is 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. The map ϕ is called the model or image of H in G.

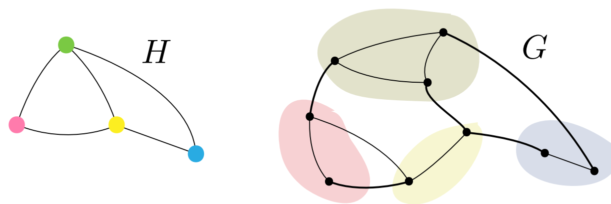

Figure 1: The graph H depicted above is a minor (at depth 1) of the graph G. The branch sets $\phi(u) \subseteq V(G)$ of vertices $u \in V(H)$ are identified by colors in the figure.

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 see that if H is a minor of G if and only if H can be obtained from G by means of the following three operations: vertex deletion, edge deletion, and edge contraction. From this 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.

In the theory of sparse graphs, we are looking at properties that are local. Therefore, excluding some graph as a minor is too restrictive for our purposes; we only need to exclude minor models that are local, in the sense that their branch sets do not connect distant parts of the graph. This is captured by the following concept of a *shallow minor*.

Definition 5. The *radius* of a connected graph G is $\mathsf{rad}(G) = \min_{u \in V(G)} \max_{v \in V(G)} \operatorname{dist}(u, v)$.

Definition 6. Let H, G be graphs and let $r \in \mathbb{N}$. The graph H is a *depth-r minor of* G, written $H \preccurlyeq_r G$, if there is a model ϕ of H in G such that the branch set $\phi(v) \subseteq G$ has radius at most r for all $v \in V(G)$. We define

$$\nabla_r(G) \coloneqq \sup \left\{ \frac{|E(H)|}{|V(H)|} \;:\; H \preccurlyeq_r G \right\}$$

and

$$\omega_r(G) \coloneqq \sup\{t \;:\; K_t \preccurlyeq_r G\}.$$

Having fixed the concept of containment in the form of shallow minors, we can proceed to defining what we mean by saying that a graph class is sparse. If one thinks about it, there are two natural definitions possible: one leads to graph classes of *bounded expansion*, and second to *nowhere dense* graph classes. These are the two central notions of the theory.

The concept of bounded expansion requires all fixed-depth minors to have bounded average degree, as explained formally below.

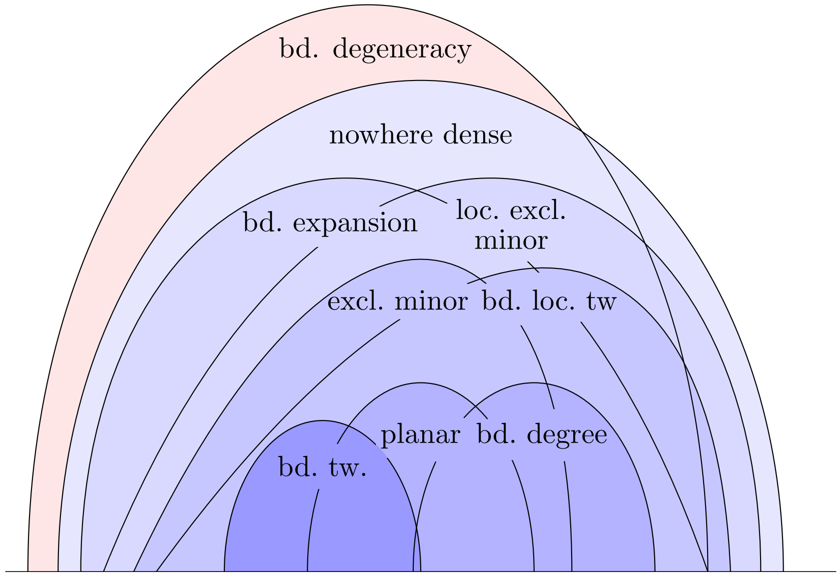

Figure 2: Illustration of the inclusions of sparse graph classes.

Definition 7. A class $\mathcal{C}$ of graphs has *bounded expansion* if there is a function $f\colon \mathbb{N} \to \mathbb{N}$ such that for all $r \in \mathbb{N}$ and for all $G \in \mathcal{C}$ we have $\nabla_r(G) \leqslant f(r)$.

Observe that the depth-0 minors of a graph are exactly its subgraphs, hence every class of bounded expansion is in particular $f(0)$-degenerate. For a graph class $\mathcal{C}$, we often write $\nabla_r(\mathcal{C}) = \sup_{G\in\mathcal{C}} \nabla_r(G)$. Note that $\mathcal{C}$ has bounded expansion iff $\nabla_r(\mathcal{C})$ is finite for every $r \in \mathbb{N}$.

The concept of nowhere denseness requires fixed-depth minors to simply not contain all cliques.

Definition 8. A class $\mathcal{C}$ of graphs is *nowhere dense* if there is a function $t\colon \mathbb{N} \to \mathbb{N}$ such that for all $r \in \mathbb{N}$ we have $\omega_r(G) \leqslant t(r)$ for all $G \in \mathcal{C}$.

Again, for a class $\mathcal{C}$ we write $\omega_r(\mathcal{C}) = \sup_{G\in\mathcal{C}} \omega_r(G)$, and $\mathcal{C}$ is nowhere dense iff $\omega_r(\mathcal{C})$ is finite for every $r \in \mathbb{N}$. Note that every graph is a subgraph of some clique, so nowhere denseness is equivalent to requiring that for every depth $r \in \mathbb{N}$ we exclude at least one graph as a depth-r 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 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 6. *Let G be a graph of maximum degree Δ. Then $\nabla_r(G) \leqslant \Delta^{r+1}$. In particular, every graph class with bounded maximum degree has bounded expansion.*

Proof. We prove that every r-shallow minor of G has maximum degree at most Δ^{r+1}. Take a depth-r 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 r from $\gamma(u)$ in $\phi(u)$. Without loss of generality $\phi(u)$ is a tree of depth at most r, 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 $r+1$ 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^{r+1} because $\overline{\phi(u)}$ is a tree of depth at most $r+1$ with maximum degree at most Δ. This concludes the proof. □

Lemma 7. *If G is an n-vertex graph that does not contain K_t as a minor, for some $t \geqslant 2$, then G has less than $2^t \cdot n$ edges. In particular, 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. We prove the claim by induction on t. For $t = 2$ the claim holds trivially, so suppose $t \geqslant 3$. For the sake of contradiction suppose there is an n-vertex graph G that has at least $2^t \cdot n$ edges but does not contain K_t as a minor. Without loss of generality let G be such a counterexample that has the minimum possible number of vertices.

Take any vertex u of G. Clearly u is not isolated in G, since otherwise the removal of u would yield a counterexample with fewer vertices. Examine any neighbor v of u and observe that u and v must have at least 2^t common neighbors. Indeed, otherwise contracting the edge uv would yield a graph $G' \preccurlyeq G$ with $n-1$ vertices and at least $2^t \cdot n - 2^t = 2^t \cdot (n-1)$ edges. Since G' a minor of G, it also does not contain K_t as a minor. This means means that G' would be a counterexample with fewer vertices than G, a contradiction.

We conclude that every neighbor of u has at least 2^t neighbors in common with u. Hence, if we denote by A the set of neighbors of u, then the graph $G[A]$ has minimum degree at least 2^t, so in particular it has at least $2^{t-1}|A|$ edges. On the other hand, $G[A]$ cannot contain K_{t-1} as a minor, because extending a minor model of K_{t-1} in $G[A]$ with a branch set consisting only of u would yield a minor model of K_t in G. We conclude that $G[A]$ is a nonempty graph with at least $2^{t-1} \cdot |A|$ edges and not containing K_{t-1} as a minor; this is a contradiction with the induction assumption.

For the "in particular" part, observe that if $\mathcal{C}$ excludes H as a minor, then it also excludes K_t as a minor where $t = |V(H)|$. Further, every minor H of a graph from $\mathcal{C}$, no matter at what depth, still excludes K_t as a minor, so the ratio between the number of edges and the number of vertices in H is at most 2^t. □

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

Definition 9. The *girth* of a graph G, denoted $\mathsf{girth}(G)$, is the length of a shortest cycle in G.

The following standard result can obtained by the probabilistic method; we will not prove it.

Lemma 8. *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 $\frac{k|V(G)|}{100}$ edges.*

Lemma 9. *The class $\mathcal{C} = \{G : \Delta(G) \leqslant \mathsf{girth}(G)\}$ is nowhere dense and has unbounded expansion.*

Proof. We first prove that the class is nowhere dense. Assume that $K_t \preccurlyeq_r G$ for some r, t and $G \in \mathcal{C}$, and fix a depth-r minor model of K_t in G. Then all vertices contained in the minor model are contained in the $(3r+1)$-neighborhood of some vertex of G. If now $G \in \mathcal{C}$ is a graph with $\mathsf{girth}(G) \leqslant 6r+3$, then also $\Delta(G) \leqslant 6r+3$, which implies that $t \leqslant 1+(6r+3)^{r+1}$ by Lemma 6. On the other hand, if $\mathsf{girth}(G) > 6r+3$, then the $(3r+1)$-neighborhood of every vertex is acyclic, which implies $t \leqslant 2$. This shows that $\omega_r(G) \leqslant 1+(6r+3)^{r+1}$.

Second, to see that $\mathcal{C}$ has unbounded expansion, note that Lemma 8 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 10. *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 $\mathsf{rad}(\phi_J(u)) \leqslant b$, and then we take $w \in V(\phi_H(v))$ witnessing that $\mathsf{rad}(\phi_H(v)) \leqslant a$, 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. □

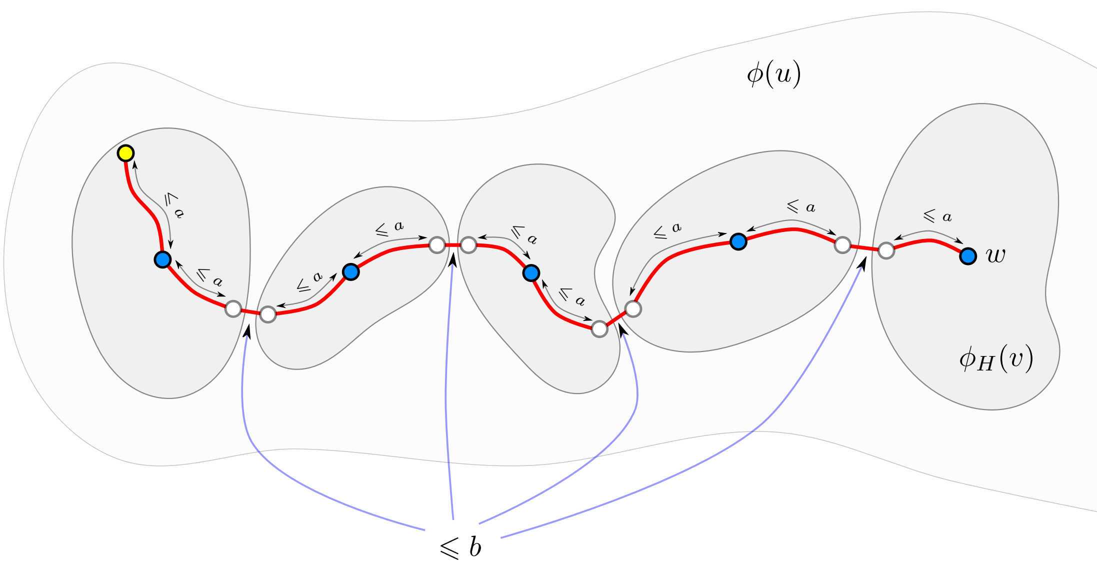

Figure 3: Proof that $\mathsf{rad}(\phi(u)) \leqslant 2ab+a+b$ for each $u \in V(J)$. Blue vertices witness that corresponding branch sets in ϕ_H have radius at most a. The red path witnesses that an arbitrary vertex of $\phi(u)$ (depicted in yellow) is at distance at most $2ab+a+b$ from w within this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

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

Definition 10. For a graph class $\mathcal{C}$ and $r \in \mathbb{N}$, we define the *depth-r reduct* of $\mathcal{C}$ as:

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

Corollary 11. *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,

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

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

3 Topological minors

In the definitions introduced the previous section we relied on the minor order as the underlying notion of topological containment for graphs. One can study also another concept of topological containment, called topological minors, which are defined as follows.

Definition 11. A graph H is a topological minor of G, written $H \preccurlyeq^{\text{top}} G$, if there is a map ϕ 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 $\phi(e)$ in G such that

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

The graph H is a *topological depth-r minor of* G, written $H \preccurlyeq^{\text{top}}_r G$, if there is a model ϕ of H in G such that the paths $\phi(e)$ have length at most $2r + 1$ for all $e \in E(H)$. We define

$$\widetilde{\nabla}_r(G) \coloneqq \sup \left\{ \frac{|E(H)|}{|V(H)|} \ : \ H \preccurlyeq^{\text{top}}_r G \right\}$$

and

$$\widetilde{\omega}_r(G) \coloneqq \sup \left\{ t \ : \ K_t \preccurlyeq^{\text{top}}_r G \right\}.$$

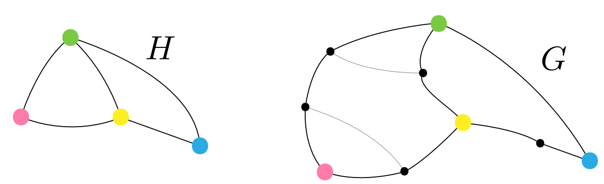

Figure 4: The graph H depicted above is a topological minor (at depth 1) of the graph G.

Observe that if H is a topological minor of G, then it is also a minor of G, and the same also holds for topological depth-r minors and depth-r minors. Again, if J is a topological minor of H

and H is a topological minor of G, then J is a topological minor of G; thus, topological minor containment is a partial order on graphs. It is, however, quite different from the minor order: a graph of maximum degree 3 may contain K_5 as a minor, but will never contain any graph with a vertex of degree 4 as a topological minor. However, it turns out that if we restrict attention to bounded-depth minors and bounded-depth topological minors, for the purpose of defining sparsity, then these notions become surprisingly similar.

3.1 Nowhere denseness via topological minors

As we observed, every topological depth-r minor is also a depth-r minor. Conversely, as we will show next, if a class $\mathcal{C}$ contains arbitrarily large complete graphs as depth-r minors, then it also contains arbitrarily large complete graphs as topological depth-s minors, for a slightly larger radius s. More precisely, we will show the following lemma.

Lemma 12. *Suppose $\mathcal{C}$ is a graph class such that for some $r \in \mathbb{N}$, $\mathcal{C}$ admits all cliques as depth-r minors; that is, every clique is a depth-r minor of some graph from $\mathcal{C}$. Then $\mathcal{C}$ admits all cliques as topological depth-$(3r+1)$ minors.*

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

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

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-r minor, there is some $G \in \mathcal{C}$ for which we can find a depth-r minor model ϕ of K_s in G. We shall prove that G also admits K_t as a topological depth-r 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.

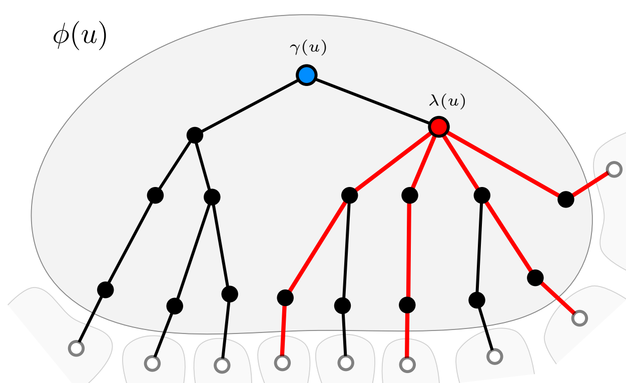

Figure 5: Choosing $\lambda(u)$ and $\mathcal{P}(u)$ (the latter depicted in red).

More precisely, similarly as in Lemma 6, 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 r 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 $r+1$ 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^{r+1}, and the depth of $\overline{\phi(u)}$ is at most $r+1$, 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-$(3r+1)$ 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 6):

- 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 r, hence there is a path of length at most $2r$ 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 $2r$ 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 $\lambda(v)$.

Observe that $P(uv)$ is a path of length at most $6r+3$ 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-$(3r+1)$ minor model of K_t in G. $\square$

Let us note that the proof of Lemma 12 actually yields a slightly stronger conclusion that describes the dependencies between parameters.

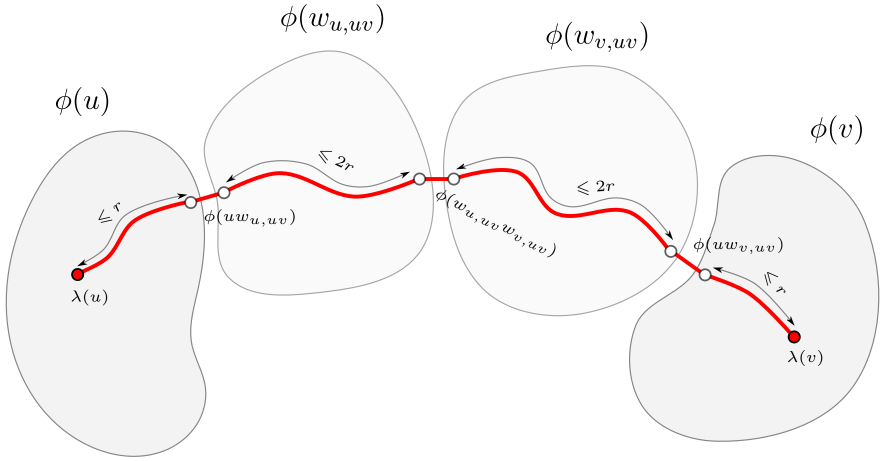

Figure 6: Construction of the path $P(uv)$ (in red).

Corollary 13. *For every graph G and $r \in \mathbb{N}$, it holds that*

$$\widetilde{\omega}_r(G) \leqslant \omega_r(G) \leqslant 2 + (\widetilde{\omega}_{3r+1}(G))^{2r+2}.$$

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

Lemma 12 implies that we can define nowhere dense classes also by excluding cliques as bounded-depth topological minors.

Corollary 14. *A class $\mathcal{C}$ of graphs is nowhere dense if and only if there is a function $t\colon \mathbb{N} \to \mathbb{N}$ such that for all $r \in \mathbb{N}$ we have $\widetilde{\omega}_r(G) \leqslant t(r)$ for all $G \in \mathcal{C}$.*

3.2 Bounded expansion via topological minors

As we have just observed, replacing shallow minors with shallow topological minors in the definition of nowhere denseness yields exactly the same notion. It turns out that the same happens also for the notion of bounded expansion. Precisely, since every depth-r shallow topological minor is also a depth-r shallow minor, we have $\widetilde{\nabla}_r(G) \leqslant \nabla_r(G)$ for every graph G, but $\nabla_r(G)$ is also bounded from above by a function of $\widetilde{\nabla}_r(G)$. This is made precise in the following lemma.

Lemma 15. *For every graph G and $r \in \mathbb{N}$, we have*

$$\widetilde{\nabla}_r(G) \leqslant \nabla_r(G) \leqslant 2^{r^2+3r+3} \cdot (\lceil \widetilde{\nabla}_r(G) \rceil)^{(r+2)^2}.$$

Corollary 16. *A class $\mathcal{C}$ of graphs has bounded expansion if and only if there is a function $f\colon \mathbb{N} \to \mathbb{N}$ such that for all $r \in \mathbb{N}$ and all $G \in \mathcal{C}$ we have $\widetilde{\nabla}_r(G) \leqslant f(r)$.*

The proof of Lemma 15 is beyond the scope of this course, due to its technicality, and was not included in the lectures. However, its conclusion, in particular Corollary 16, is important for the theory. For completeness and for interested readers, we include a proof of Lemma 15 below.

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

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. □

Let us now recall some intuition from the proof of Lemma 12. Suppose ϕ is a depth-r 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 r) in ϕ if $\overline{\phi(u)}$ is a star with every edge subdivided at most r times (i.e., replaced by a path of length at most $r + 1$), and the leaves of $\overline{\phi(u)}$ are exactly its legs. Observe that if all vertices of H are spiders of depth at most r in ϕ, then ϕ can be trivially turned into a depth-r topological minor model H in G.

Therefore, intuitively our goal in the proof of Lemma 15 is the following: starting with a depth-r minor model ϕ of some really dense graph H in G, transform it into a depth-r 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 12, 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 7 for an illustration.

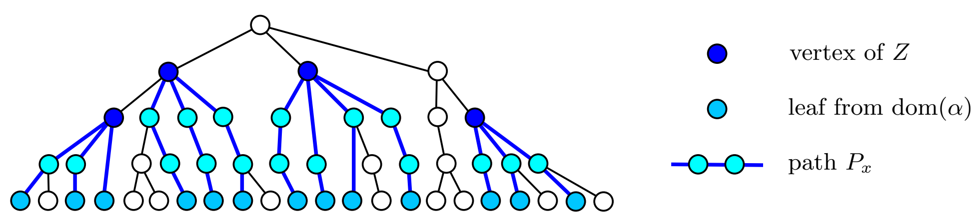

Figure 7: An example tree of depth 4 and a possible outcome of applying Lemma 18 on it, for $d = 2$. Each leaf $x \in \mathrm{dom}(\alpha)$ is mapped by α to the vertex of Z to which the path P_x leads.

Lemma 18. *Suppose T is a rooted tree of depth at most $r + 1$ and with ℓ leaves, and let d 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 d leaves that are all descendants of z.*
- *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 $\lfloor \frac{\ell - 1}{d^{r+1}} \rfloor$.*

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

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

From Lemma 18 we deduce the following statement, where we replace many branch sets with spiders in one shot.

Lemma 19. *Suppose H is a bipartite graph, say with bipartition (X, Y), whose edge density is more than $(2d)^{r+2}$, where d is a positive integer. Let ϕ be a depth-r minor model of H in some graph G. Then there exists a bipartite graph H' with bipartition (X', Y), and a depth-r minor model ϕ' of H' in G, with the following properties:*

- *the edge density of H' is more than d;*
- *each $u \in X'$ is a spider of depth at most r in ϕ'; and*
- *for each $v \in Y$, we have $\phi'(v) = \phi(v)$.*

Proof. Similarly as in the proof of Lemma 12, we may assume that each branch set $\phi(u)$ is a tree of depth at most r 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 $r+1$ 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 18 to the tree $\overline{\phi(u)}$, for parameter $2d$. 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 $a \in \alpha_u^{-1}(w)$, take the vertex $v \in Y$ such that $a \in V(\phi(v))$, and make w adjacent to v in H'. We may construct a depth-r 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 $a \in \alpha_u^{-1}(w)$, the path in $\overline{\phi(u)}$ between w and a 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 r in ϕ'.

We are left with making sure that the edge density of H' is more than d. For a vertex $u \in X$, let $d_H(u)$ be the degree of u in H; then $|E(H)| = \sum_{u \in X} d_H(u)$. By Lemma 18, we have that each vertex of X' has degree more than $2d$ in H', thus

$$|E(H')| > 2d|X'|. \tag{3.1}$$

On the other hand, by Lemma 18 again, we have that each vertex $u \in X$ gives rise to at least $\lfloor \frac{d_H(u)-1}{(2d)^{r+1}} \rfloor \geqslant \frac{d_H(u)}{(2d)^{r+1}} - 1$ edges in H'. Therefore, we have

$$|E(H')| \geqslant \sum_{u \in X} \left(\frac{d_H(u)}{(2d)^{r+1}} - 1 \right) = \frac{|E(H)|}{(2d)^{r+1}} - |X| \geqslant \frac{(2d)^{r+2}}{(2d)^{r+1}}(|X| + |Y|) - |X| \geqslant 2d|Y|. \tag{3.2}$$

By combining (3.1) and (3.2) we infer that

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

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

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

Proof of Lemma 15. Let $d \coloneqq \lceil \widetilde{\nabla}_r(G) \rceil$, and let function f be defined as $f(x) = (2x)^{r+2}$; then we need to prove that $\nabla_r(G) \leqslant 2 \cdot f(f(d))$. For the sake of contradiction, suppose there is a depth-r minor H_0 of G with edge density more than $2 \cdot f(f(d))$. By Lemma 17, H_0 contains a bipartite subgraph H_1 with edge density more than $f(f(d))$; obviously H_1 is still a depth-r minor of G. Let (X, Y) be a bipartition of H_1, and let ϕ_1 be a depth-r minor model of H_1 in G.

We apply Lemma 19 to H_1 and its model ϕ_1 in G, with parameter $f(d)$. This yields a bipartite graph H_2 with bipartition (X', Y), and a depth-r minor ϕ_2 of H_2 in G, such that: the edge density of H_2 is more than $f(d)$, 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 19, this time to H_2 and its model ϕ_2 in G, and with parameter d, 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-r minor ϕ_3 of H_3 in G, such that: the edge density of H_3 is more than d, 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 $r+1$, 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 r in ϕ_3, so as we argued before, the model ϕ_3 in fact witnesses that H_3 is a depth-r topological minor of G. However, the edge density of H_3 is larger than $d = \lceil \widetilde{\nabla}_r(G) \rceil$, a contradiction. □

4 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 20. *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 $n \geqslant N$ vertices, we have that G has less than $n^{1+\varepsilon}$ edges.*

Observe that Theorem 20 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 20 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 20 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 21. *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 $|V(G')| \geqslant \sqrt{n}$.*

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 22 (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*

$$\mathbb{P}(|S - \mu| > \delta\mu) \leqslant 2\exp\left(-\frac{\delta^2\mu}{3}\right).$$

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

Lemma 23. *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 $n \geqslant M(\varepsilon)$ vertices with minimum degree at least n^{ε}. Then either G contains K_t as a depth-1 minor for some $t \geqslant \log n$, or we can find a depth-1 minor G' of G with $n' \geqslant n^{1-\varepsilon}$ 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

$$\mathbb{P}(n^{1-\varepsilon}\log n \leqslant |A| \leqslant 3n^{1-\varepsilon}\log n) \geqslant 1 - 2\exp\left(-\frac{n^{1-\varepsilon}\log n}{6}\right). \tag{4.1}$$

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:

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

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

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

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

$$\mathbb{P}\left(|B| \geqslant \frac{n}{2}\right) \geqslant \frac{2}{3}. \tag{4.2}$$

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 $t \geqslant \log n$ 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 (4.1) and (4.2) we argued that, provided $M(\varepsilon)$ is chosen large enough, each of the following events happens with probability at least $\frac{2}{3}$:

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

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:

$$\frac{|E(G')|}{|V(G')|} \geqslant \frac{|B|}{|A|} \geqslant \frac{\frac{n}{2}}{3n^{1-\varepsilon}\log n} = \frac{n^{\varepsilon}}{6\log n}. \tag{4.3}$$

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

$$n \geqslant \left(\frac{n'}{3\log n}\right)^{\frac{1}{1-\varepsilon}}. \tag{4.4}$$

Putting (4.3) and (4.4) together, and using the assumption that $\varepsilon \leqslant \frac{2}{3}$, we infer that

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

Now observe that

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

Hence, we have

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

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

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

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

With all the tools prepared, we can wrap up the proof of Theorem 20.

Proof of Theorem 20. Observe that since the class $\mathcal{C}\nabla r$ is also nowhere dense (Corollary 11), 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 21 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 23 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 $t \geqslant \log n_{i-1}$, 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 $n_i \geqslant (n_{i-1})^{\frac{1-\varepsilon}{2}} \geqslant (n_{i-1})^{\frac{1}{6}}$ 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 $t \geqslant \log n_i$, because the other outcome would be a graph with a super-quadratic number of edges.

The above procedure, if applicable, terminates after at most $k := \lceil \frac{1}{\varepsilon^2} \rceil$ iterations, for in each iteration we have that ε_i is larger by at least $\varepsilon_{i-1}^2 \geqslant \varepsilon^2$ 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 10 we get that G_i is a depth-d_i minor of G for $d_i := \frac{3^i-1}{2}$. 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 23 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 $M(\varepsilon_i) \leqslant M(\varepsilon)$ 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 23 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 $N(\varepsilon)$. $\square$