Chapter 3: Uniform quasi-wideness

Compilation date: December 3, 2017

1 Introduction

In the previous chapters we have characterized nowhere dense classes by means of sizes of cliques that we find as depth-r minors. In this chapter we study a dual characterization in terms of *r-independence*. Recall that a vertex subset $A \subseteq V(G)$ in a graph G is r-independent if any two its elements $a \neq b \in A$ are at distance greater than r. In degenerate graphs we are able to find large independent subsets, however, this is no longer true if we ask for larger values of r. Consider for example a star S, i.e., a tree of depth 1. No two vertices of S are at distance greater than 1. However, if we were allowed to delete a bounded number of elements, we could delete the center of the star and in the resulting graph we are left with r-independent elements, whatever value for r we choose. We will formalize this concept of deleting a few elements to find large r-independent sets by introducing the notion of *uniform quasi-wideness.* We will then show that the new concept is equivalent to nowhere denseness and show how to use it in an algorithmic context.

2 Uniform wideness

Let us first consider the simpler concept of *uniform wideness*, which will help to understand uniform quasi-wideness.

Definition 1. A class of graph $\mathcal{C}$ is called *uniformly wide* if there is a function $N\colon \mathbb{N}\times\mathbb{N}\to\mathbb{N}$ such that for all $m, r \in N$, $G \in \mathcal{C}$ and $A \subseteq V(G)$ with $|A| \geqslant N(m,r)$ there exists $B \subseteq A$ with $|B| \geqslant m$ such that B is r-independent.

In other words, a class of graphs is uniformly wide if for every value of r, in every *huge* set A we still find a *large* r-independent set. The appropriate definition of *huge* and *large* depends on the value of r we care for. It is not difficult to see that uniformly wide classes are very simple classes, as the next theorem shows.

Theorem 1. *A class $\mathcal{C}$ of graphs is uniformly wide if and only if $\mathcal{C}$ is a class of bounded degree, i.e., there is a number d such that the maximum degree $\Delta(G)$ of every $G \in \mathcal{C}$ is bounded by d.*

Proof. Assume first that $\mathcal{C}$ is uniformly wide. By definition there is a function $N\colon \mathbb{N}\times\mathbb{N}\to\mathbb{N}$ such that for all $m, r \in N$, $G \in \mathcal{C}$, and $A \subseteq V(G)$ with $|A| \geqslant N(m,r)$ there exists $B \subseteq A$ with $|B| \geqslant m$ such that B is r-independent. We claim that for each $G \in \mathcal{C}$, the maximum degree of G is smaller than $N(2,2)$. Take any vertex v of G and let $A = N(v)$. Then all vertices of A are pairwise at distance at most 2, and hence there is no set $B \subseteq A$ of size 2 that would be 2-independent. We infer that $|N(v)| = |A| < N(2,2)$, and this must hold for every vertex v of G.

Conversely, assume that $\mathcal{C}$ has bounded degree, and let d be an integer such that the maximum degree $\Delta(G)$ of every $G \in \mathcal{C}$ is bounded by d. Define $N(m,r) = m \cdot (d+1)^r$. We claim that $\mathcal{C}$ is uniformly wide with function N. To see this, let $G \in \mathcal{C}$ and $A \subseteq V(G)$ with $|A| \geqslant N(m,r)$ for some numbers m and r. We can now greedily pick elements from A to the set B as follows. Choose an arbitrary vertex $v \in A$ and put it into the set B, then remove all elements at distance at most r from v from the set A. As every vertex has degree at most d, we remove at most $(d+1)^r$ vertices from A in each step, which after m steps gives us the desired set B. $\square$

3 Uniform quasi-wideness

We now slightly change the definition of wideness to allow the deletion of a small number of vertices.

Definition 2. A class of graph $\mathcal{C}$ is called *uniformly quasi-wide* if there are functions $N\colon \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ and $s\colon \mathbb{N} \to \mathbb{N}$ such that for all $m, r \in N$, $G \in \mathcal{C}$ and $A \subseteq V(G)$ with $|A| \geqslant N(m,r)$ there exists $S \subseteq V(G)$ with $|S| \leqslant s(r)$ and $B \subseteq A - S$ with $|B| \geqslant m$ such that B is r-independent in $G - S$.

In other words, a class of graphs is uniformly quasi-wide if for every value of r and for every *huge* set A we can delete a *very small* number of vertices such that we find a *large* r-independent subset of A in $G - S$. Again the appropriate definitions of *huge*, *large* and *very small* depend on the value of r we care for; this is governed by functions N and s that are sometimes called the *margins*.

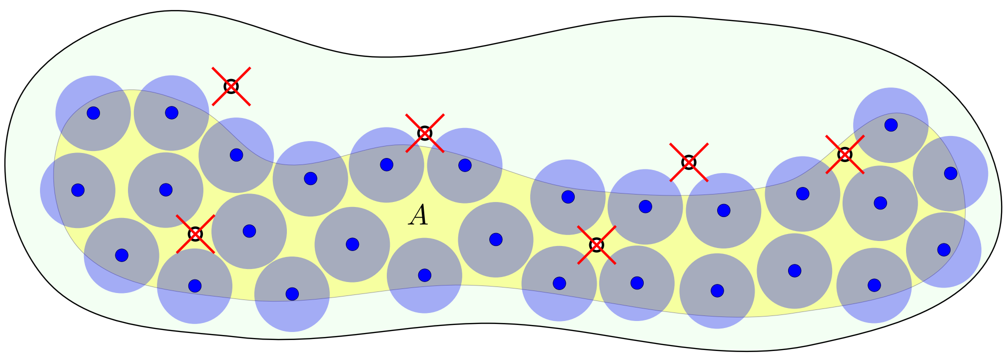

Figure 1: Definition of uniform quasi-wideness. In a huge set A (yellow) we may find a large subset B (blue) that is r-independent after removing a small subset of vertices S (crossed out), whose size depends only on the radius r. In case r is even, r-independence of B in $G - S$ is equivalent to saying that balls of radius $r/2$ around vertices of B in $G - S$ are disjoint, as depicted in the figure.

The rest of this section is devoted to the surprising fact that uniform quasi-wide classes are exactly the nowhere dense classes.

Theorem 2. *A class $\mathcal{C}$ of graphs is uniformly quasi-wide if and only if it is nowhere dense.*

We split the proof into two lemmas. The direction from left to right is easy to prove.

Lemma 3. *If $\mathcal{C}$ is uniformly quasi-wide, then $\mathcal{C}$ is nowhere dense.*

Proof. We prove the contrapositive. Supposing $\mathcal{C}$ is somewhere dense, there exists r such that every complete graph K_t is a depth-r topological minor of some $G \in \mathcal{C}$. Fix any functions $N\colon \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ and $s\colon \mathbb{N} \to \mathbb{N}$. Choose $G \in \mathcal{C}$ such that $K \coloneqq K_t \preccurlyeq_r^{\text{top}} G$, where $t \coloneqq N(2 \cdot s(2r+1)+2, 2r+1)$; fix a depth-$r$ topological minor model ϕ of K in G. Let A be the set $\{\phi(v) \colon v \in V(K)\}$. Consider any vertex subset $S \subseteq V(G)$ of size at most $s \coloneqq s(2r+1)$, and let $A' \subseteq A$ be constructed from

A as follows: whenever $\phi(u) \in S$ for some $u \in V(K)$, remove $\phi(u)$ from A', and whenever an internal vertex of $\phi(uv)$ belongs to S for some $uv \in E(K)$, remove both u and v from A'. Thus, when constructing A' from A we remove at most $2s$ vertices, meaning $|A - A'| \leqslant 2s$. Now observe that every two vertices of A' are at distance at most $2r+1$ in $G - S$, because the path between them in the model ϕ was left untouched by the removal of S. Hence, any subset $B \subseteq A - S$ that is $(2r+1)$-independent in $G - S$ contains at most $2s+1$ vertices: at most $2s$ vertices of $A - A'$ and at most 1 vertex of A'. Since $|A| = t = N(2s+2, 2r+1)$ and S was chosen arbitrarily, this witnesses that N and s are not valid wideness functions for $\mathcal{C}$. As the choice for these functions was arbitrary, we conclude that $\mathcal{C}$ is not uniformly quasi-wide. □

The other direction is much harder to prove.

Lemma 4. *If $\mathcal{C}$ is nowhere dense, then $\mathcal{C}$ is uniformly quasi-wide.*

We will split the proof into several more lemmas. In the following, fix a function $t : \mathbb{N} \to \mathbb{N}$ and a graph G such that for all $r \in \mathbb{N}$ we have $K_{t(r)} \not\preccurlyeq_r G$. Also fix a large $A \subseteq V(G)$. Our strategy is to inductively construct sequences

$$G = G_0 \supseteq G_1 \supseteq \ldots \supseteq G_r \qquad \text{and} \qquad S_1, S_2, \ldots, S_r,$$

where G_i are graphs and S_i are vertex sets such that for all $i \in \{1, \ldots, r\}$, we have

1. $G_i = G_{i-1} - S_i$ and
2. $S_i \subseteq V(G_{i-1})$.

Moreover, we will find a sequences

$$A = A_0 \supseteq A_1 \supseteq \ldots \supseteq A_r \qquad \text{and} \qquad m_0 \geqslant m_1 \geqslant \ldots \geqslant m_r = m,$$

where A_i are vertex sets and m_i are integers, such that for all $i \in \{1, \ldots, r\}$,

1. $A_i \subseteq A$ is i-independent in $G_{i-1} - S_i$, $A_i \cap S_i = \emptyset$,
2. $|A_i| \geqslant m_i$, and
3. $S_i = \emptyset$ if i is odd and $|S_i| < t(i/2)$ if i is even.

We will then return the set A_r of size m which is r-independent in $G_r = G - S$, where $S = \bigcup_{1 \leqslant i \leqslant r} S_r$. The construction will be applicable provided initially the invariant $|A_0| \geqslant m_0$ holds, hence we will set $N(m, r)$ simply as the obtained m_0.

The following two lemmas will imply that it suffices to consider only the cases $i = 1$ and $i = 2$. Their proofs are immediate.

Lemma 5. *Let A be a $2j$-independent set in G. Let $H \preccurlyeq_j G$ be the depth-j minor of G obtained by contracting the disjoint j-neighborhoods $N_j^G[v]$ for $v \in A$ to single vertices. The vertex of H resulting from contracting $N_j^G[v]$ will be identified with the original vertex v of G, thus via this identification A is both a subset of vertices of G and a subset of vertices of H. Then any subset of A is $(2j+1)$-independent in G if and only if it is 1-independent in H.*

Lemma 6. *Let A be a $2j+1$-independent set in G. Let $H \preccurlyeq_j G$ be the depth-j minor of G obtained by contracting the disjoint j-neighborhoods $N_j^G[v]$ for $v \in A$ to single vertices. The vertex of H resulting from contracting $N_j^G[v]$ will be identified with the original vertex v of G, thus via this identification A is both a subset of vertices of G and a subset of vertices of H. Then, for any $S \subseteq V(H) - A = V(G) - \bigcup_{v \in A} N_j(v)$, it holds that a subset of A is $2j+2$-independent in $G - S$ if and only if it is 2-independent in $H - S$.*

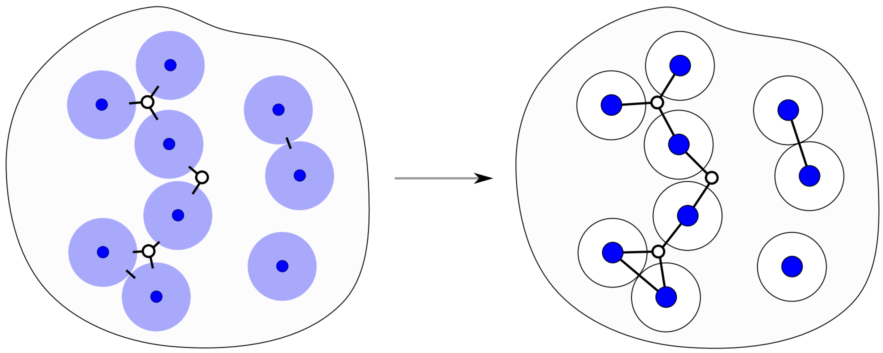

Figure 2: Reduction to cases $i = 1$ and $i = 2$ in Lemmas 5 and 6: contracting balls of radius j around vertices of a $2j$-independent set turns $(2j+1)$-independence into 1-independence, and $(2j+2)$-independence into 2-independence.

The construction will be done iteratively for $i = 1, 2, \ldots, r$, where in step i we wish to construct A_i and S_i by examining the graph G_{i-1} and the $(i-1)$-independent set A_{i-1} in it. The whole procedure will work as follows, assuming that for the cases $i = 1$ and $i = 2$ we have already given the construction. Starting with the set $A = A_0$, using the case $i = 1$ we will find a large independent subset $A_1 \subseteq A$ without deleting any vertices, hence $S_1 = \emptyset$, as claimed. Now that A_1 is 1-independent, using the case $i = 2$ we will find a small set S_2 and a large subset A_2 of A_1 which is 2-independent in the graph $G_2 = G_1 - S_2$ (hence, the set is 2-independent after deleting S_2, just as required in the definition of uniform quasi-wideness). As A_2 is 2-independent in G_2, we can contract the disjoint 1-neighborhoods of elements of A_2 (identifying contracted vertices with elements of A_2, as in Lemmas 5 and 6), thus obtaining a depth-1 minor $H \preccurlyeq_1 G_2$. Using the case $i = 1$ again, we find a large subset $A_3 \subseteq B_2$ that is 1-independent in H and apply Lemma 5 to conclude that it is, in fact, 3-independent in G_3. We continue with this set in the graph G_3. Again, we contract the disjoint 1-neighborhoods of elements of A_3 (identifying contracted vertices with elements of A_3), thus obtaining a depth-1 minor $H \preccurlyeq_1 G_3$. Using case $i = 2$ in H we find a small set $S_4 \subseteq V(H) - A_3 = V(G_3) - \bigcup_{a \in A_3} N_1^{G_3}[a]$ and a large set $A_4 \subseteq A_3$ which is 2-independent in $H - S_4$. We apply Lemma 6 to conclude that A_4 that A_4 is 4-independent in $G_4 = G_3 - S_4$; the lemma is applicable since $S_4 \subseteq V(H) - A_4$. We continue this argumentation for r steps to construct the graphs G_i and sets A_1 and S_i with the desired properties.

It remains to show cases $i = 1$ and $i = 2$: how to construct A_1 out of A_0, and how to construct A_2, S_2 out of A_1. One of the main ingredients for this is Ramsey's Theorem.

Theorem 7. *Let $a, b \in \mathbb{N}$. Then there exists a number $R(a,b)$ such that for every coloring of the edges of a complete graph on $R(a,b)$ vertices with colors red and blue we will either find a clique on a vertices whose edges are all blue or a clique on b vertices whose edges are all red.*

Proof. We prove by induction on $a+b$ that it suffices to take $R(a,b) = R(a-1,b) + R(a,b-1)$. Clearly, we for all $n \in \mathbb{N}$ we may take $R(n,1) = R(1,n) = 1$; then, it is easy to see by induction that the above recurrence will yield

$$R(a,b) \leqslant \binom{a+b-2}{a-1}.$$

Assume that we have established the bounds for $R(a-1,b)$ and $R(a,b-1)$ and consider a complete graph K on $R(a-1,b) + R(a,b-1)$ vertices whose edges are colored red and blue. Pick a vertex v and partition the remaining vertices into two sets A and B, such that for every vertex $w \in A$ the edge vw is blue and for every vertex $w \in B$ if the edge vw is red. We have $|A| + |B| + 1 = R(a-1,b) + R(a,b-1)$, and hence either $|A| \geqslant R(a-1,b)$ or $|B| \geqslant R(a,b-1)$. In the first case, by induction we know that M contains either a clique on $a-1$ vertices with all edges blue, or a clique on b vertices with all edges red. In the latter subcase we are immediately done, and in the former case we may add v to this clique to obtain a clique on a vertices will all edges blue. The second case is analogous. This finishes the proof of the theorem. □

We may now give the construction for $i = 1$. Recall that $t(0)$ is such that $K_{t(0)} \not\preccurlyeq_0 G$, that is, $K_{t(0)}$ is not a subgraph of G. Suppose we are given a set A of size $|A| \geqslant m_0 := \binom{m_1+t(0)-2}{t(0)-1}$, where m_1 is the target size of a 1-independent set we are interested in. By Ramsey's theorem, in $G[A]$ we may either find a clique of size $t(0)$ or an independent set of size m_1. The former case, however, cannot happen since $K_{t(0)}$ is not a subgraph of G. So we obtain an independent set of size m_1, as promised. To lifting this to the case of $i = 2j+1$ using Lemma 6, as explained before, we will apply this argument to a graph that is a j-shallow minor of the original graph, hence we will need to set $m_{i-1} := \binom{m_i+t(i/2)-2}{t(i/2)-1}$. Note again that $S_i = \emptyset$ in this case.

The case $i = 2$, which will lift to the induction step for even i, is much harder. In this case we assume that in our graph G we have already found a huge 1-independent set A_1, and we want to find a large 2-independent set in it, possibly after removing some small set of vertices S_1 that is disjoint from A_1; we will have that $|S_1| < t := t(1)$. Denote by D the set of all neighbors of vertices of A_1 and consider the graph G' defined as the subgraph of G on the vertex set $A_1 \cup D$ where we preserve only edges with one endpoint in A_1 and second in D. Clearly G' is bipartite, with bipartition $A_1 \uplus D$. Since A_1 is independent in G, it is easy to see that any subset of A_1 is 2-independent in G if and only if it is 2-independent in G', so we may focus on G'. We will use the following extension of Ramsey's Theorem for a finite number of colours.

Theorem 8. *Let $n_1, \ldots, n_k \in \mathbb{N}$. There exists a number $R(n_1, \ldots, n_k)$ such that for every coloring of the edges of a complete graph on $R(n_1, \ldots, R_k)$ vertices with k different colors $c_1, \ldots, c_k$ we will find for some $1 \leqslant i \leqslant k$ a clique on n_i vertices all of whose edges are colored with color c_i.*

The theorem can easily be proved by induction on the number of colors, using the two-color case. We will apply the theorem to prove the following lemma.

Lemma 9. *Let G be a bipartite graph with partitions A and B. Let $m,t,d \in \mathbb{N}$. If $|A| \geqslant R(t, \ldots, t, m)$, where t is repeated $\binom{d-1}{2}$ times, then at least one of the following assertions holds.*

(a) A contains a set $A' \subseteq A$ of size m such that no two vertices of A' have a common neighbor,

(b) in G there is a 1-subdivision of K_t with all principal vertices contained in A, or

(c) B contains a vertex of degree at least d.

Let us motivate the statement of the lemma by examining its application to the graph G' we discussed before. The lemma says that provided A_1 is sufficiently large, we will either find a 2-independent set, which is exactly what we are looking for, or a 1-subdivision of K_t, which should not happen in a nowhere dense class, or we may find a vertex v in D that has degree at least d. We will add this vertex to the set S_1 of vertices to delete and inductively continue with the set $A_1' = N(v) \cap A_1$. We will apply the lemma again to the bipartite graph induced by A_1' and its neighborhood and with v deleted (with a smaller value d'), again, giving us a 2-independent subset (and we are done), a 1-subdivision of K_t (which is again not possible by assumption) or another vertex of high degree. We apply the lemma again and again, always on the subset of A_1 induced by the neighborhood of the high degree vertex, which eventually, if the initial value of d was chosen large enough, gives us a complete bipartite graph $K_{t,t}$. This however, is not possible, as $K_{t,t}$ contains K_t itself as a depth-1 minor. We conclude that before we could apply the lemma t times, we must have found a large 2-independent set. Precise argument will follow, but now we give a proof of Lemma 9.

Proof of Lemma 9. Assume that B does not contain a vertex of degree at least d (otherwise we conclude that the third assertion holds). Enumerate the vertices of B as $b_1, \ldots, b_n$. Let K be the complete graph with vertex set A whose edges we will color with $\binom{d-1}{2} + 1$ colors. We initially consider all edges as colorless. Now we consider the vertices $b_1, \ldots, b_n$ in increasing order. In each step i, $1 \leqslant i \leqslant n$, we consider the set $X = N(b_i) \cap A$ and color each edge uv for $u, v \in X$ which has not previously received a color with an integer between 1 and $\binom{d-1}{2}$ such that no two edges that are colored in this step get the same color. This is possible as the degree of b_i is assumed to be at most $d-1$. Finally, we color all edges which do not have received a color in the steps $1, \ldots, n$ with color $\binom{d-1}{2} + 1$.

As $|A| \geqslant R(t, \ldots, t, m)$, in the resulting colored graph we can either find a clique of size t whose edges are all colored with one of the colors $1, \ldots, \binom{d-1}{2}$, or a clique of size m whose edges are all colored with color $\binom{d-1}{2}+1$. In the latter case, the vertices of the clique define a subset $A' \subseteq A$ such that no two vertices in A' have a common neighbor, as stated in the first assertion of the lemma. In the former case, all edges of the monochromatic clique have been added at different steps in the construction, as all edges receive different colors in each individual step. Hence the edges have been colored in $\binom{t}{2}$ different steps, and each edge can be associated with a different vertex b_i that caused coloring of exactly this edge. Hence we find a 1-subdivision of K_t with all principal vertices in A. This is the second assertion of the lemma. $\square$

As we discussed, the idea is to apply Lemma 9 to the graph G' defined earlier not once, but t times. For convenience, we write $R_d(t, m)$ for $R(t, \ldots, t, m)$ where the first argument is repeated $\binom{d-1}{2}$ times. Then let

$$R^\star(t, m, d, s) := \begin{cases} t & \text{if s=0} \\ R_k(t, m) & \text{if } s \geqslant 1 \text{, where } k = R^\star(t, m, d, s-1). \end{cases}$$

The next lemma explains the iterative application of Lemma 9.

Lemma 10. *Let G be a bipartite graph with partitions A and B. If $|A| \geqslant R^\star(t, m, t, t)$, then at least one of the following assertions holds.*

(a) A contains a set $A' \subseteq A$ of size m and B contains a set S of size lass than t such that no two vertices of A' have a common neighbor outside of S,

(b) in G there is a 1-subdivision of K_t with all principal vertices contained in A,

(c) in G there is a complete bipartite subgraph $K_{t,t}$.

Proof. We will iteratively find vertices $s_1, s_2, \ldots$ and subsets $A = A_0 \supseteq A_1 \supseteq A_2 \supseteq \ldots$ with $|A_i| \geqslant R^\star(t, m, t, t-i)$ and $A_i \subseteq N(s_1) \cap N(s_2) \cap \ldots N(s_i)$. Suppose $s_1, \ldots, s_i$ and A_i are already defined for some $i < t$. Then apply Lemma 9 to the graph $G[A_i \cup B_i]$, where $B_i = B - \{s_1, \ldots, s_i\}$. This application yields either objects witnessing the satisfaction of the first or the second assertion (for $S = \{s_1, \ldots, s_i\}$), or gives a vertex s_{i+1} that has at least $R^\star(t, m, t, t-i-1)$ neighbors in A_i. In the former two cases we may stop the iteration, and in the latter case we may define $A_{i+1} \coloneqq A_i \cap N(s_{i+1})$ and proceed. Finally, observe that if $s_1, \ldots, s_t$ and A_t (with $|A_t| \geqslant R^\star(t, m, t, 0) = t$) have been constructed, then $\{s_1, \ldots, s_t\} \cup A_t$ is a complete bipartite graph in G, so the third assertion holds. □

Using Lemma 10 we may solve directly the case $i = 2$ of the main construction. Assuming that $|A_1| \geqslant m_1 \coloneqq R^\star(t, m_2, t, t)$, where m_2 is the requested size of a 2-independent set after this step, apply Lemma 10 to the bipartite graph G' we defined. This application cannot yield either a 1-subdivision of K_t or a $K_{t,t}$ subgraph of G', since both these graphs contain K_t as a 1-shallow minor, which is excluded since we assumed $K_t \not\preccurlyeq_1 G$. The last conclusion — a set $S_1 \subseteq D$ with $|S_1| < t$ together with a subset $A_2 \subseteq A_1$ with $|A_2| \geqslant m_2$ that is 2-independent in $G' - S$ — is exactly what we were looking for.

As we discussed earlier, the case $i = 2$ presented above lifts to all even i by applying Lemma 6. More precisely, we apply case $i = 2$ to the graph H obtained from the contractions and the i-independent set A_i in it. Observe that in this setting, to exclude assertions (b) and (c) in Lemma 10 it suffices to take $t \coloneqq t(i/2)$. Indeed, if in the contracted graph H we find either a 1-subdivision of K_t with all principal vertices in A_i, or a $K_{t,t}$ with one side contained in A_i and second outside of A_i, then in the graph before contractions, both of these would yield a depth-$(i/2)$ minor model of K_t, a contradiction.

To summarize, we put

$$m_{i-1} \coloneqq \begin{cases} \binom{m_i + t(i/2) - 2}{t(i/2) - 1} & \text{if } i \text{ is odd;} \\ R^\star(t(i/2), m_i, t(i/2), t(i/2)) & \text{if } i \text{ is even.} \end{cases}$$

By requesting $m_r \coloneqq m$, this gives a value of $N(m, r) \coloneqq m_0$ for which the whole construction can be performed. This concludes the proof of Lemma 4, which was the missing part of the proof of Theorem 2.

The bound on $N(m, r)$ given by the proof might be considered infeasible for practical applications. To prevent the reader from these diffuse depreciating sentiments, we will in a later chapter revisit the proof and provide much improved bounds.

Finally, let us note an effective version of the theorem. It follows by examining all the components of the proof and observing that they can be turned into polynomial-time algorithms computing respective objects.

Theorem 11. *Let $\mathcal{C}$ be nowhere dense. Then there are functions $N : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ and $s : \mathbb{N} \to \mathbb{N}$ such that on input $G \in \mathcal{C}$, $m, r \in \mathbb{N}$ and $A \subseteq V(G)$ we can compute in polynomial time a set $S \subseteq V(G)$ with $|S| \leqslant s(r)$ and $B \subseteq A - S$ with $|B| \geqslant m$ such that B is r-independent in $G - S$.*

4 The splitter game

We finally provide a very intuitive game characterization of nowhere denseness.

Definition 3. Let G be a graph and let $\ell, m, r > 0$. The (ℓ, m, r)-*splitter game* on G is played by two players, "Connector" and "Splitter", as follows. We let $G_0 := G$. In round i of the game, Connector chooses a vertex $v_i \in V(G_{i-1})$. Then Splitter picks a subset $W_i \subseteq N_r^{G_{i-1}}(v_i)$ of size at most m. We let $G_i := G_{i-1}[N_r^{G_{i-1}}(v_i)] - W_i$. Splitter wins if $G_i = \emptyset$. Otherwise the game continues at G_i. If Splitter has not won after ℓ rounds, then Connector wins.

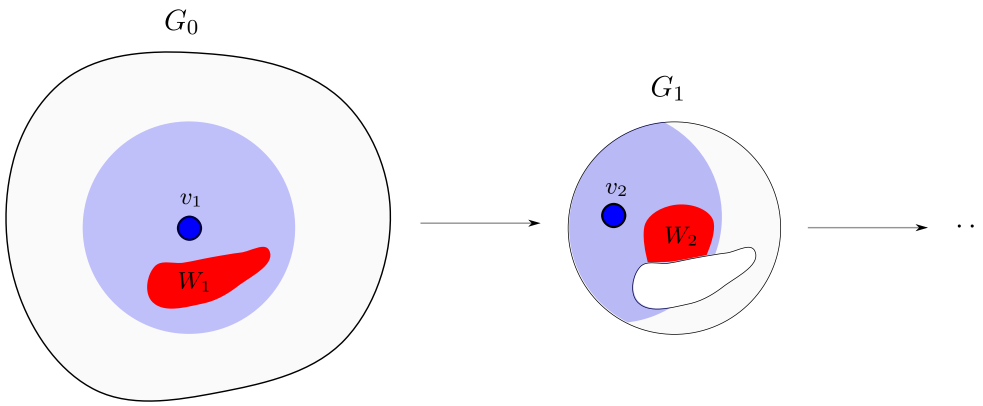

Figure 3: First two rounds of the splitter game. In the ith round, Connector first picks a vertex v_i and the arena gets restricted to the r-neighborhood of v_i. Then the Splitter removes a set W_i of at most m vertices from the arena. The goal of Splitter is to obtain an empty graph within ℓ rounds, the goal of Connector is to prevent this.

In the splitter game, a *strategy* for Splitter is, well, what one expects it to be. Formally, it is a function f that maps every partial play $(v_1, W_1, \ldots, v_s, W_s, v_{s+1})$ to a new move $W_{s+1} \subseteq N_r^{G_s}(v_{s+1})$ of Splitter. Similarly for Connector. A strategy f is a *winning strategy* for Splitter in the (ℓ, m, r)-splitter game on G if Splitter wins every play in which he follows the strategy f. If Splitter has a winning strategy, we say that he *wins* the (ℓ, m, r)-splitter game on G.

We first show that nowhere denseness guarantees that Splitter wins the splitter game.

Theorem 12. *Let $\mathcal{C}$ be a nowhere dense class of graphs. Then for every $r \in \mathbb{N}$ there are $\ell \in \mathbb{N}$ and $m \in \mathbb{N}$, such that for every $G \in \mathcal{C}$, Splitter wins the (ℓ, m, r)-splitter game on G.*

Proof. As $\mathcal{C}$ is nowhere dense, it is also uniformly quasi-wide, and let s and N be the functions witnessing this. Fix $r \in \mathbb{N}$ and let $\ell := N(r, 2s(r) + 1)$ and $m := \ell \cdot (r + 1)$. Note that both ℓ and m only depend on $\mathcal{C}$ and r. We claim that for any $G \in \mathcal{C}$, Splitter wins the (ℓ, m, r)-splitter game on G; for this, we present a suitable winning strategy.

Let $G \in \mathcal{C}$ be a graph. In the (ℓ, m, r)-splitter game on G, Splitter uses the following strategy. In the first round, if Connector chooses $v_1 \in V(G_0)$, where $G_0 := G$, then Splitter chooses $W_1 :=$

$\{v_1\}$. Now let $i \geqslant 1$ and suppose that $v_1, \ldots, v_{i-1}, G_1, \ldots, G_{i-1}, W_1, \ldots, W_{i-1}$ have already been defined. Suppose Connector chooses $v_i \in V(G_{i-1})$. We define W_i as follows. For each $j < i$, choose a path $P_{j,i}$ in G_{j-1} of length at most r connecting v_j and v_i. Such a path must exist as $v_i \in V(G_i) \subseteq V(G_j) \subseteq N_r^{G_{j-1}}[v_j]$. We let $W_i \coloneqq \bigcup_{1 \leqslant j < i} V(P_{j,i}) \cap N_r^{G_{i-1}}(v_i)$; in other words, the move W_i of Splitter consists of all vertices of path $P_{j,i}$ that are still in the arena. Note that $|W_i| \leqslant (i-1) \cdot (r+1) \leqslant m$, as the paths have length at most r and hence consist of $r+1$ vertices. It remains to be shown is that the length of any such play is bounded by ℓ.

Assume toward a contradiction that Connector may survive for ℓ rounds. Let $(v_1, \ldots, v_\ell, G_1, \ldots, G_\ell, W_1, \ldots, W_\ell)$ be a play witnessing this, where the moves of the Splitter are according to the presented strategy. As $\ell \geqslant N(r, 2s(r)+2)$, for $A \coloneqq \{v_1, \ldots, v_\ell\}$ there is a set $S \subseteq V(G)$ with $|S| \leqslant s(r)$, such that $A - S$ contains subset B of size $2s(r)+2$ that is r-independent in $G - S$. Suppose $B = \{v_{i_1}, \ldots, v_{i_{2s(r)+2}}\}$ with $i_1 < \ldots < i_{2s(r)+2}$; for brevity we write $w_j \coloneqq v_{i_j}$.

We now consider the pairs (w_{2j-1}, w_{2j}) for $1 \leqslant j \leqslant s(r)+1$. By construction, $Q_j \coloneqq P_{i_{2j-1}, i_{2j}}$ is a path of length at most r from w_{2j-1} to w_{2j} in $G_{i_{2j-1}-1}$. We now observe that paths Q_j, for $j \in \{1, \ldots, s(r)+1\}$, are pairwise disjoint. Indeed, if $1 \leqslant j < j' \leqslant s(r)+1$, then the whole path Q_j was removed by the Splitter in round i_{2j} (formally, $V(Q_j) \cap V(G_{i_{2j-1}}) \subseteq W_{i_{2j}}$), hence it is entirely disjoint with the vertex set of the graph $G_{i_{2j'-1}-1}$ due to $j' > j$. On the other hand, since $Q_{j'}$ is entirely contained in the graph $G_{i_{2j'-1}-1}$ by definition, indeed Q_j and $Q_{j'}$ are disjoint. Now, since S contains at most $s(r)$ vertices, some path Q_j has to be entirely disjoint with S. However, this means that w_{2j-1} and w_{2j} do not belong to S and are at distance at most r in $G - S$. This contradicts the assumption that B is r-indendent in $G - S$ and finished the proof. □

We observe that also the converse of Theorem 12 holds and hence the splitter game provides another characterization of nowhere dense classes of graphs.

Theorem 13. *Let $\mathcal{C}$ be a class of graphs. If for every $r \in \mathbb{N}$ there are $\ell, m \in \mathbb{N}$ such that for every graph $G \in \mathcal{C}$, Splitter wins the (ℓ, m, r)-splitter game, then $\mathcal{C}$ is nowhere dense.*

Proof. We prove the contrapositive. Suppose $\mathcal{C}$ is somewhere dense, hence $\mathcal{C}$ admits all complete graphs as depth-r minors, for some fixed depth $r \in \mathbb{N}$. Then we claim that for all $\ell, m \in \mathbb{N}$ there is a graph $G \in \mathcal{C}$ such that Connector wins the $(\ell, m, 4r+1)$-splitter game on G.

Fix $\ell, m \in \mathbb{N}$. We choose $G \in \mathcal{C}$ such that G contains the complete graph $K \coloneqq K_{\ell m+1}$ as a depth-$(4r+1)$ minor; let ϕ be a minor model of K in G witnessing this. Connector uses the following strategy to win the $(\ell, m, 4r+1)$-splitter game. First, Connector chooses any vertex from the branch set (under ϕ) of any vertex of K. The $(4r+1)$-neighborhood of this vertex contains the whole branch sets of all vertices of K. Splitter removes any m vertices. We actually allow him to remove the complete branch sets of ϕ containing all m vertices he chose. In round 2 we may thus assume that we still find the complete graph $K_{(\ell-1)m+1}$ as a depth-r minor of the current arena. By continuing to play in this way until, after round ℓ the arena still contains some vertices and the Connector wins. □

5 Algorithmic applications: r-Dominating Set

As an algorithmic application of uniform quasi-wideness we show how to use it in the design of efficient algorithms for the (parameterized) r-Dominating Set problem. Recall that in a graph G, a subset of vertices D is *r-dominates* a subset of vertices A if every vertex of A is at distance at most r from a vertex of D. Further, D is an *r-dominating set* of G if D is r-dominates the whole vertex set. By $\mathrm{dom}_r(G, A)$ we denote the smallest size of an r-dominator of A in G, and $\mathrm{dom}_r(G)$ — the r-domination number of G — is the smallest size of an r-dominating set in G. We shall consider the following decision problem called r-Dominating Set: given a graph G and parameter k, is it true that $\mathrm{dom}_r(G) \leqslant k$? In general, our goal will be two-fold: (a) to design an efficient algorithm for this problem assuming r and k are small, and (b) to reduce the size of the instance at hand to a function of r and k only. In the research terminology, point (a) is to design an efficient *parameterized* algorithm for the problem, and point (b) is to design a *kernelization* procedure.

The main idea is to reduce the number of *dominatees*, that is, the number of those vertices whose domination is essential. To formalize the notion of being essential for domination, we give the following definition.

Definition 4. Let G be a graph and $k \in \mathbb{N}$. A set $Z \subseteq V(G)$ is called an *r-domination core for parameter k* if every set $D \subseteq V(G)$ of size at most k which r-dominates Z also r-dominates $V(G)$.

Fix a nowhere dense class $\mathcal{C}$ of graphs and let $N(m, r)$ and $s(r)$ be the functions characterizing $\mathcal{C}$ as uniformly quasi-wide according to Theorem 11. Fix positive integers r and k and let $s := s(2r)$. The next, slightly surprising lemma shows that in an r-domination core that is too large one can always find an *irrelevant dominatee* that can be safely removed.

Lemma 14. *Suppose $G \in \mathcal{C}$ and let $Z \subseteq V(G)$ be a vertex subset satisfying*

$$|Z| > N\big((k+2)(r+1)^s, 2r\big).$$

Then we can compute in polynomial time a vertex $w \in Z$ such that for any set $D \subseteq V(G)$ with $|D| \leqslant k$, the following equivalence holds:

$$D \text{ } r\text{-dominates } Z \text{ if and only if } D \text{ } r\text{-dominates } Z - \{w\}.$$

Proof. By Theorem 11 we can find in polynomial time sets $S \subseteq V(G)$ and $B \subseteq Z - S$ such that $|S| \leqslant s$, $|B| \geqslant (k+2)(r+1)^s$ and B is $2r$-independent in $G - S$. For each $v \in B$, compute $\pi_r[v, S]$, the r-distance profile of v on S; recall that $\pi_r[v, S]$ is a function from S to $\{0, 1, \ldots, r, \infty\}$ such that for $a \in S$ we put $\pi_r[v, S](a) = \mathrm{dist}(v, a)$ if this distance is at most r, and $\pi_r[v, S](a) = \infty$ otherwise. Clearly, we can compute these distance profiles in polynomial time. Note that there are at most $(r+1)^s$ different r-distance profiles on S. Since $|B| \geqslant (k+2)(r+1)^s$, there are $k+2$ elements $b_1, \ldots, b_{k+2} \in B$ which have the same distance profile. Now we choose $w := b_1$ and show that for any set $D \subseteq V(G)$ with $|D| \leqslant k$, D r-dominates Z if and only if D r-dominates $Z - \{b_1\}$.

The direction from left to right is obvious. Now, suppose D r-dominates $Z - \{b_1\}$. Consider the sets $W_i := N_r^{G-S}[b_i]$ for $i \in \{2, \ldots, k+2\}$. Since B is $2r$-independent in $G - S$, the sets W_i are pairwise disjoint. Since there are $k+1$ of these sets, at least one of them, say W_j, does not contain any element of D. However, since $b_j \in Z - \{b_1\}$ and D r-dominates $Z - \{b_1\}$, there is a path of length at most r from some element $x \in D$ to b_j. This path must, therefore, go through an element of S. Since b_1 and b_j have the same r-distance profiles on S, we conclude that there is also a path of length at most r from x to b_1 and therefore D r-dominates Z. □

An immediate corollary of Lemma 14 is that we can always find a small domination core. Simply start with $Z = V(G)$, which is an r-domination core vacuously, and apply the above procedure to remove an irrelevant dominatee from the r-domination core Z until $|Z| \leqslant N\big((k+2)(r+1)^s, 2r\big)$.

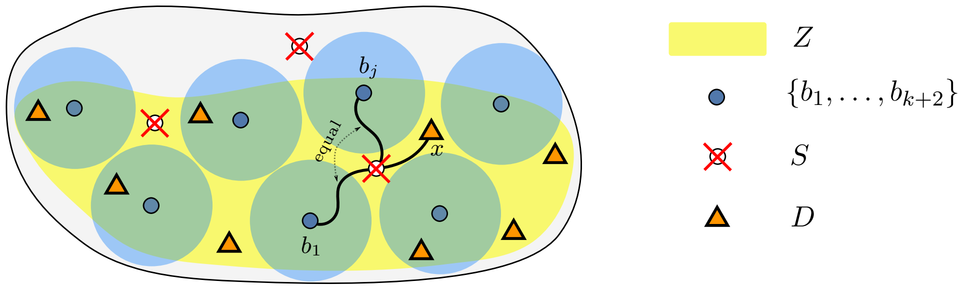

Figure 4: Situation in the proof of Lemma 14.

Corollary 15. *There is an algorithm running in polynomial time that given a graph $G \in \mathcal{C}$ computes an r-domination core of G for parameter k of size at most $N\big((k+2)(r+1)^s, 2r\big)$.*

Having a small domination core is already sufficient to design an efficient parameterized algorithm for the problem, via standard dynamic programming on subsets.

Lemma 16. *Given a graph G with n vertices and m edges, vertex subset Z, and radius $r \in \mathbb{N}$, one can compute $\mathrm{dom}_r(G, Z)$ in time $2^{|Z|} \cdot |Z|^{\mathcal{O}(1)} \cdot (n+m)$.*

Proof. First, we compute the distance $\mathrm{dist}(u, z)$ for all $u \in V(G)$ and $z \in Z$; for this it suffices to run a breadth-first search from each vertex of Z, which takes time $\mathcal{O}(|Z|(n+m))$. Thus, we may assume that with each vertex $u \in V(G)$ we store the set $L_u := N_G^r[u] \cap Z$ of vertices of Z r-dominated by u. Let us enumerate $V(G)$ as $u_1, u_2, \ldots, u_n$. We shall compute the following dynamic programming table: for $0 \leqslant i \leqslant n$ and $X \subseteq Z$, we define

$$\mathsf{D}[i, X] := \text{smallest size of a subset of } \{u_1, \ldots, u_i\} \text{ that } r\text{-dominates } X.$$

It is clear that $\mathsf{D}[0, \emptyset] = 0$ and $\mathsf{D}[0, X] = \infty$ for $X \neq \emptyset$. Moreover, the following recurrence is easy to verify:

$$\mathsf{D}[i, X] = \min\left(\mathsf{D}[i-1, X]\,,\, 1 + \mathsf{D}[i-1, X - L_{u_i}]\right).$$

The arguments of the minimum respectively correspond to the cases when u_i is not taken or taken to the constructed solution. Using this recurrence it is straightforward to compute all the values of $\mathsf{D}[i, X]$ by iterating through all i from 1 to n. Then the value $\mathsf{D}[n, Z]$ is equal to $\mathrm{dom}_r(G, Z)$. $\square$

Corollary 17. *For any nowhere dense class $\mathcal{C}$, the r-Dominating Set problem on an n-vertex graph from $\mathcal{C}$ can be solved in time $f(r, k) \cdot n^c$ for some function f and a universal constant c, indepndent of $\mathcal{C}$.*

Proof. Using Corollary 15, in polynomial time compute an r-domination core Z in G of size at most $g(r, k)$, for some function g depending only on $\mathcal{C}$. By definition of an r-domination core we have that $\mathrm{dom}_r(G) \leqslant k$ if and only if $\mathrm{dom}_r(G, Z) \leqslant k$. Then use Lemma 16 to compute $\mathrm{dom}_r(G, Z)$ in time $2^{g(r,k)} \cdot g(r, k)^{\mathcal{O}(1)} \cdot n^2$, and check whether it is not larger than k. $\square$

We now move to the second algorithmic corollary, namely reducing the instance size to a function of k and r. For this, we need to reduce the number of *dominators*, that is, the number of vertices that shall be used to dominate other vertices. Obviously, only vertices at distance at most r to a vertex from the r-domination core are relevant. Furthermore, if there are two vertices $v, v' \in V(G)$ with $N_r[v] \cap Z = N_r[v'] \cap Z$, it suffices to keep one of u and v as a representative.

Theorem 18. *Suppose $\mathcal{C}$ is a nowhere dense class of graphs. Then there exists a polynomial time algorithm that on input $G \in \mathcal{C}$ and $r, k \in \mathbb{N}$ computes an induced subgraph $H \subseteq G$ and a vertex subset $Z \subseteq V(H)$ such that* $\mathrm{dom}_r(G) \leqslant k$ *if and only if* $\mathrm{dom}_r(H, Z) \leqslant k$*. Furthermore, H has size bounded by a function of k and r only.*

Proof. Using the algorithm of Corollary 15, we first compute an r-domination core Z in G of size at most $N\big((k+2)(r+1)^s, 2r\big)$.

Now, for every vertex $v \in V(G)$ we compute the set $L_v \coloneqq N_r^G[v] \cap Z$. Consider two vertices $v, v' \in V(G)$ *equivalent* if $L_v = L_{v'}$. Clearly, the number of equivalence classes of this relation is at most $2^{|Z|}$, hence let A be any set of at most $2^{|Z|}$ vertices containing one element from each equivalence class. Finally, construct a set W as follows: start with putting $A \cup Z$ into W and then, for every pair of vertices $u, v \in A \cup Z$, if $\mathrm{dist}_G(u, v) \leqslant r$ then add the vertices of any path of length at most r between u and v to W. Clearly the size of W computed in this manner is bounded by a function of k and r only, and we are left with verifying that $H \coloneqq G[W]$ and Z satisfy the asserted property: $\mathrm{dom}_r(G) \leqslant k$ iff $\mathrm{dom}_r(H, Z) \leqslant k$. Since Z is an r-domination core in G, we have $\mathrm{dom}_r(G) \leqslant k$ iff $\mathrm{dom}_r(G, Z) \leqslant k$. Hence, it suffices to prove that $\mathrm{dom}_r(G, Z) = \mathrm{dom}_r(H, Z)$.

In one direction, if $D \subseteq V(H)$ r-dominates Z in H, then D also r-dominates Z in G, because H is an induced subgraph of G. This proves that $\mathrm{dom}_r(G, Z) \leqslant \mathrm{dom}_r(H, Z)$.

In the other direction, take any $D \subseteq V(G)$ that r-dominates Z in G. For each $x \in D$, some vertex x' that is equivalent to x has been included in A. Let $D' \coloneqq \{x' \colon x \in D\}$; clearly $|D'| \leqslant |D|$ and $D' \subseteq A \subseteq W$. It is now straightforward to see that D' r-dominates Z in H, since for each $x \in D$ the corresponding vertex $x' \in D'$ r-dominates exactly the same vertices of Z in H as x r-dominated in G. This is because we explicitly added to H a path of length at most r between x' and every vertex of Z that was r-dominated by x' in G. This proves that $\mathrm{dom}_r(G, Z) \geqslant \mathrm{dom}_r(H, Z)$, ergo $\mathrm{dom}_r(G, Z) = \mathrm{dom}_r(H, Z)$ and we are done. □

Again, the size of the "kernel" (H, Z) provided by Theorem 18 may be impractical. However, in the next section we will refine this result by showing that, for bounded expansion classes, the kernel size may be reduced to *linear* in k. For nowhere dense classes, the size may be reduced to almost linear, but this is a more difficult result (but not much more difficult).

6 Linear kernelization for r-Dominating Set

We will now improve the conclusion of Theorem 18 by proving that in bounded expansion classes one may compute a kernel for the r-Dominating Set problem whose size is *linear* in k, the target size of the r-dominating set. More precisely, our goal in this section is the following theorem.

Theorem 19. *Let $\mathcal{C}$ be a class of bounded expansion and let $r \in \mathbb{N}$ be fixed. Then there exists a polynomial-time algorithm that, given a graph $G \in \mathcal{C}$ and $k \in \mathbb{N}$, either correctly concludes that* $\mathrm{dom}_r(G) > k$*, or computes an induced subgraph H of G and a subset of its vertices $Z \subseteq V(H)$ such that* $\mathrm{dom}_r(G) = \mathrm{dom}_r(H, Z)$ *and $|V(H)| \leqslant ck$ for some constant c depending only on $\mathcal{C}$ and r.*

The proof of Theorem 19 will essentially involve all the tools that we have seen so far during this course. The main ingredients will be (a) the approximation algorithm for r-Dominating Set, (b) neighborhood complexity, and (c) uniform quasi-wideness. We will also need a number of technical lemmas that were given during the tutorials. We recall and formally prove them first.

6.1 Toolbox

We will work with the concepts of *projections* and *projection profiles*, which appeared at least twice during tutorials: once in the beginning in the context of the so-called *Projection Closure lemma*, and then in the context of neighborhood complexity. The setting is as follows. Let G be a graph and let A be a subset of vertices of G. A path P connecting some $u \in V(G) - A$ with some $a \in A$ is called A-*avoiding* if all vertices traversed by P, apart from the endpoint a, do not belong to A. For $r \in \mathbb{N}$, the r-*projection* of a vertex $u \in V(G) - A$ onto A, denoted $M_r(u, A)$, is the set of all vertices of A reachable from u by an A-avoiding path of length at most r. The r-*projection profile* of a vertex $u \in V(G) - A$ on A is the function $\mu_r[u, A] \colon A \to \{1, \ldots, r, \infty\}$ defined as follows: for $a \in A$, the value $\mu_r[u, A](a)$ is the length of a shortest A-avoiding path connecting u and a, or ∞ if this length is larger than r. Thus $M_r(u, A) = (\mu_r[u, A])^{-1}(\{1, \ldots, r\})$. A function $f \colon A \to \{1, \ldots, r, \infty\}$ is *realized* as an r-projection profile on A if there exists $u \in V(G) - A$ such that $f = \mu_r[u, A]$.

During the previous lectures we have seen that in a bounded expansion class, the number of different r-distance profiles on a set A is bounded linearly in the size of A. During the tutorials we have lifted this argument to r-projection profiles.

Lemma 20. *Let $\mathcal{C}$ be a class of bounded expansion and let $r \in \mathbb{N}$. There exists a constant c, depending only on $\mathcal{C}$ and r, such that for every $G \in \mathcal{C}$ and nonempty $A \subseteq V(G)$, the number of different functions from A to $\{1, \ldots, r, \infty\}$ realized as r-projection profiles on A is at most $c|A|$.*

Proof. Take any $G \in \mathcal{C}$ and $A \subseteq V(G)$. Construct G' from G by removing all edges with both endpoints in A and subdividing $2r$ times every edge with exactly one endpoint in A. Vertices of G are naturally identified with the corresponding vertices of G'. It is straightforward to see that for every vertex $u \in V(G) - A$, its r-projection profile in G, $\mu_r^G[u, A]$, is equal to its $3r$-distance profile in G', $\pi_{3r}^{G'}[u, A]$, where all finite values are shifted by $2r$. Observe also that if $\mathcal{C}$ has bounded expansion, then the class $\mathcal{C}'$ comprising graphs obtained from graphs from $\mathcal{C}$ by subdividing every edge an arbitrary number of times also has bounded expansion; this is because every depth-r minor of a graph from $\mathcal{C}'$ can be obtained from a depth-r minor of a graph from $\mathcal{C}$ by adding vertices of degree 2 only. Thus $G' \in \mathcal{C}'$, hence the number of different $3r$-distance profiles on A in G' is bounded by $c|A|$, for some constant c depending only on $\mathcal{C}$ and r. Consequently, the number of different r-projection profiles on A in G is also bounded by $c|A|$. $\square$

Of course, Lemma 20 also shows that in a graph from class of bounded expansion the number of different r-projections on a set A is bounded linearly in $|A|$. Next, we will need the aforementioned Projection Closure Lemma. Intuitively, it says that every vertex subset can be "closed" to a set that admits r-projections only of bounded size, at the cost of blowing up the size linearly.

Lemma 21 (Projection Closure Lemma)**.** *Let $\mathcal{C}$ be a class of bounded expansion and let $r \in \mathbb{N}$. Then there exists a constant c and a polynomial-time algorithm that, given a graph $G \in \mathcal{C}$ and a vertex subset $A \subseteq V(G)$, computes a vertex subset $A' \supseteq A$ such that $|A'| \leqslant c|A|$ and $|M_r(u, A')| \leqslant c$ for all $u \in V(G) - A'$.*

Proof. Let $d = \lceil 2\nabla_{r-1}(\mathcal{C}) \rceil$. We perform an iterative procedure which maintains a graph H, initially set to G, and a subset of its vertices B, initially set to A. Every iteration of the procedure works as follows. Check whether there exists a vertex $u \in V(H) - B$ with $|M_r(u, B)| \geqslant d$. If this is not the case, terminate the procedure outputting H and B. Otherwise, select any d distinct vertices $b_1, \ldots, b_d$ from $M_r(u, B)$ together with A-avoiding paths $P_1, \ldots, P_d$ of length at most r, where P_i connects u with b_i. Modify H by contracting the subgraph $\bigcup_{i=1}^{d} P_i - \{b_i\}$ onto u, add the vertex resulting from this contraction to B, and proceed with the iteration. It is straightforward to implement this procedure in polynomial time.

Observe that during the above procedure, each contraction affects only vertices not belonging to B and the vertex resulting from the contraction is being put into B, so it does not participate in any further contractions. Moreover, each contraction is applied to a connected subgraph of radius $r-1$. It follows that throughout the procedure, we maintain the invariant that H is a depth-$(r-1)$ minor of G.

We claim that the procedure terminates after at most $|A|$ iterations. Suppose otherwise, and let H and B be the state after $|A|+1$ iterations. In particular B consists of $2|A|+1$ vertices: $|A|$ original vertices of A and $|A|+1$ vertices resulting from contractions that were added in subsequent iterations. Observe that each added vertex u, at the moment of adding it to B, brought at least d edges to $H[B]$; these are edges connecting it to $b_1, \ldots, b_d$, in the notation from the description of the procedure. It follows that after $2|A|+1$ iterations, the graph $H[B]$ contains at least $d(|A|+1)$ edges. However, this means that

$$\frac{|E(H[B])|}{|V(H[B])|} \geqslant \frac{d(|A|+1)}{2|A|+1} > \frac{d}{2} \geqslant \nabla_{r-1}(\mathcal{C}),$$

which contradicts the fact that H is a depth-$(r-1)$ minor of G.

Let then H and B be the results of the procedure, returned by it at the moment of termination. Note that each vertex $b \in B$ is either an original vertex of A or has been obtained by contracting a subgraph of G on at most $d(r-1)+1$ vertices. In the latter case, let $\phi(b)$ be the set of vertices of this subgraph (that was contracted onto b). In the former case, when $b \in A$, we let $\phi(b) = \{b\}$. Thus $|\phi(b)| \leqslant d(r-1)+1$ for all $b \in B$.

Let now $A' \coloneqq \bigcup_{b \in B} \phi(b)$. Clearly $A \subseteq A' \subseteq V(G)$ and we have

$$|A'| \leqslant (d(r-1)+1) \cdot |B| \leqslant 2(d(r-1)+1) \cdot |A|.$$

Thus, it suffices to prove that the r-projections on A' of vertices from $V(G)-A'$ have sizes bounded by a constant. Take any $u \in V(G')-A'$; observe that $V(G')-A' = V(H)-B$, so u is also a vertex of $V(H)-B$. By the condition of termination of the procedure, we have that $|M_r^H(u,B)| < d$. On the other hand, it is straightforward to see that $M_r^{G'}(u,A') \subseteq \bigcup_{b \in M_r^H(u,B)} \phi(b)$. This is because under the contractions that yield H from G, every A'-avoiding path in G connecting u with a vertex of $\phi(b)$ is mapped to a B-avoiding path in H of the same length connecting u with b. Hence,

$$|M_r^{G'}(u,A')| \leqslant (d(r-1)+1) \cdot |M_r^H(u,B)| < d(d(r-1)+1).$$

Concluding, it suffices to take $c \coloneqq \max(2(d(r-1)+1), d(d(r-1)+1))$. □

The last tool that we will use is another closure lemma. This time we are concerned about finding a small superset of a given vertex subset A that is sufficient to preserve distances, up to threshold r, between elements of A. During the tutorials we first gave a proof of this results using Projection Closure Lemma and lexicographic products. Then, in the third homework we suggested another proof using generalized coloring numbers. We now give the second proof, as it is conceptually simpler.

Lemma 22 (Shortest Paths Closure Lemma)**.** *Let $\mathcal{C}$ be a class of bounded expansion and let $r \in \mathbb{N}$. Then there exists a constant c and a polynomial-time algorithm that, given a graph $G \in \mathcal{C}$ and a vertex subset $A \subseteq V(G)$, computes a vertex subset $A' \supseteq A$ such that $|A'| \leqslant c|A|$ and whenever $\mathrm{dist}_G(u,v) \leqslant r$ for some $u, v \in A$, then also $\mathrm{dist}_{G[A']}(u,v) = \mathrm{dist}_G(u,v)$.*

Proof. As we discussed in the previous lectures, we may in polynomial time compute a vertex ordering σ of G such that $\mathrm{wcol}_r(G,\sigma) \leqslant d$ for some constant d depending only on $\mathcal{C}$ and r. Now, for every vertex $a \in A$ let $L_a \subseteq V(G)$ be constructed as follows: for every $u \in \mathrm{WReach}_r[G,\sigma,a]$, add to L_a the vertex set of an arbitrarily chosen shortest path from a to u. As each of these paths brings at most r vertices to L_a apart from a itself, we have $|L_a| \leqslant 1 + rd$.

Now define $A' \coloneqq \bigcup_{a \in A} L_a$. Clearly $|A'| \leqslant (1+rd)|A|$, so we will put $c \coloneqq 1 + rd$. It remains to prove the last assertion about A'. For this, take any $u, v \in A$ with $\mathrm{dist}_G(u,v) \leqslant r$. Obviously $\mathrm{dist}_{G[A']}(u,v) \geqslant \mathrm{dist}_G(u,v)$, since $G[A']$ is an induced subgraph of G, hence we only need to prove the converse inequality. Let P be any shortest path between u and v and let w be the vertex of P that is the smallest in σ. The prefix of P from u to w and the suffix of P from w to v witness that $w \in \mathrm{WReach}_r[G,\sigma,u] \cap \mathrm{WReach}_r[G,\sigma,v]$. Consequently, A' contains the vertex sets of some shortest paths (in G) between u and w and between u and v. Therefore, we have

$$\mathrm{dist}_{G[A']}(u,v) \leqslant \mathrm{dist}_{G[A']}(u,w) + \mathrm{dist}_{G[A']}(w,v) = \mathrm{dist}_G(u,w) + \mathrm{dist}_G(w,v) = \mathrm{dist}_G(u,v),$$

where the last equality follows from the fact that w lies on a shortest path from u to v. □

6.2 Phase 1: reducing dominatees

Similarly as in the previous section, the kernelization algorithm will proceed in two phases. First, we reduce *dominatees* — intuitively, a dominatee is a vertex that is required to be dominated, and we iteratively remove these constraints from vertices until only a linear number of dominatees is left. Second, we reduce *dominators* — vertices that may be taken to an r-dominating set. Once we have only a linear number of dominatees and dominators, all the other vertices may only serve the purpose of describing the metric in the graph. As made explicit in the Shortest Paths Closure Lemma, we only need to preserve a linear number of them to represent faithfully distances between dominators and dominatees.

In this subsection we give the first, main phase of the algorithm: reduction of dominatees. From now on we fix a class of bounded expansion $\mathcal{C}$, graph $G \in \mathcal{C}$, radius $r \in \mathbb{N}$, and target dominating set size $k \in \mathbb{N}$.

The concept of dominatees is made formal via the notion of a *domination core*, as in the previous section. Unfortunately, due to technical reasons that will become clear later, we will need a slightly different definition of a core, which we shall call a *strict core* for clarity. For convenience, we introduce the following notation. For a vertex subset $Z \subseteq V(G)$, a *dominator* of Z is any subset of vertices D that r-dominates Z; recall that the size of a smallest dominator of Z is denoted by $\mathrm{dom}_r(G,Z)$. A dominator of Z is *optimal* if it has the minimum possible size $\mathrm{dom}_r(G,Z)$.

Definition 5. A *strict core* in G is any subset of vertices $Z \subseteq V(G)$ such that every optimal dominator of Z is also an r-dominating set of G.

Observe that the above definition does not say anything about suboptimal dominators of Z; they might not be r-dominating sets of G. It is only required that any set that r-dominates Z optimally is automatically an r-dominating set in G. The following observation is straightforward.

Lemma 23. *If Z is a strict core in G then* $\mathrm{dom}_r(G, Z) = \mathrm{dom}_r(G)$.

Proof. Clearly $\mathrm{dom}_r(G, Z) \leqslant \mathrm{dom}_r(G)$, because in the former we only need to r-dominate a subset of vertices. On the other hand, since every optimal dominator of Z is also an r-dominating set in G, we have $\mathrm{dom}_r(G, Z) \geqslant \mathrm{dom}_r(G)$. □

Throghout the rest of this section we will present an *irrelevant dominatee rule*: in a large strict core Z we may always find, in polynomial time, a vertex that can be safely removed from Z. This is made formal in the following lemma.

Lemma 24. *There is a constant c, depending only on $\mathcal{C}$ and r, such that given a vertex subset Z with $|Z| > ck$ and a guarantee that Z is a strict core, one can in polynomial time either correctly conclude that* $\mathrm{dom}_r(G) > k$ *or find a vertex $z \in Z$ such that $Z - \{z\}$ is also a strict core.*

As before, we may start with $Z = V(G)$, which is trivially a strict core, and remove vertices from Z one by one using the algorithm of Lemma 24 until the size of the strict core at hand is at most ck. Thus, Lemma 24 immediately implies the following.

Lemma 25. *There is a constant c, depending only on $\mathcal{C}$ and r, such that one can in polynomial time either correctly conclude that* $\mathrm{dom}_r(G) > k$ *or compute a strict core Z in G of size at most ck.*

We now proceed with the proof of Lemma 24, so we assume that we are given a subset of vertices Z with $|Z| > ck$ and a guarantee that Z is a strict core. The value of the constant c will be determined in the course of the proof, and it will be a function of multiple other constants depending on $\mathcal{C}$ and r, given by various tools that we will use. Each of them will be denoted by c with a subscript succinctly describing the origin of the constant. We describe how to combinatorially find a vertex z that can be excluded from the strict core Z; the proof can be trivially turned into a polynomial-time algorithm.

The first move is to apply the approximation algorithm for r-Dominating Set. Recall that in the fifth lecture we gave a polynomial-time algorithm that, given a graph G from a fixed bounded expansion class $\mathcal{C}$ and radius $r \in \mathbb{N}$, computes an r-dominating set in G of size at most $c_{\mathrm{apx}} \cdot \mathrm{dom}_r(G)$, for some constant c_{apx} depending only on $\mathcal{C}$ and r. Apply this algorithm to G and let D_{apx} be the obtained r-dominating set in G. If $|D_{\mathrm{apx}}| > c_{\mathrm{apx}}k$, then we are sure that $\mathrm{dom}_r(G) > k$ and we may terminate the algorithm and report this result. Thus, from now on we may assume that $|D_{\mathrm{apx}}| \leqslant c_{\mathrm{apx}}k$.

Next, we apply the Projection Closure Lemma to D_{apx} for parameter $3r$, yielding its superset $A \supseteq D_{\mathrm{apx}}$ with the following properties: $|A| \leqslant c_{\mathrm{prj}}|D_{\mathrm{apx}}|$ and every $3r$-projection on A of a vertex outside of A has size at most c_{prj}, for some constant c_{prj} depending only on $\mathcal{C}$ and r. Observe that thus $|A| \leqslant c_{\mathrm{prj}}c_{\mathrm{apx}}k$ and A is also an r-dominating set in G, as it contains D_{apx}.

Now let us examine the set $Z' := Z - A$. We classify the vertices of Z' according to their $3r$-projections on A. More precisely, we define following equivalence relation: for $u, v \in Z'$, we put

$$u \sim v \qquad \text{if and only if} \qquad \mu_{3r}(u, A) = \mu_{3r}(v, A).$$

By Lemma 20, the number of equivalence classes of $\sim$ is bounded by $c_{\mathrm{nei}} \cdot |A|$ for some constant c_{nei} depending only on $\mathcal{C}$ and r. Note that this number is upper bounded by $c_{\mathrm{nei}}c_{\mathrm{prj}}c_{\mathrm{apx}}k$.

Now suppose that we have $c = c_{\mathrm{nei}}c_{\mathrm{prj}}c_{\mathrm{apx}}c' + c_{\mathrm{prj}}c_{\mathrm{apx}}$, for some constant c' that we will determine later. Since $|Z| > ck$ and $|A| \leqslant c_{\mathrm{prj}}c_{\mathrm{apx}}k$, we have $|Z'| > c_{\mathrm{nei}}c_{\mathrm{prj}}c_{\mathrm{apx}}c' \cdot k$. Since $\sim$ has at most

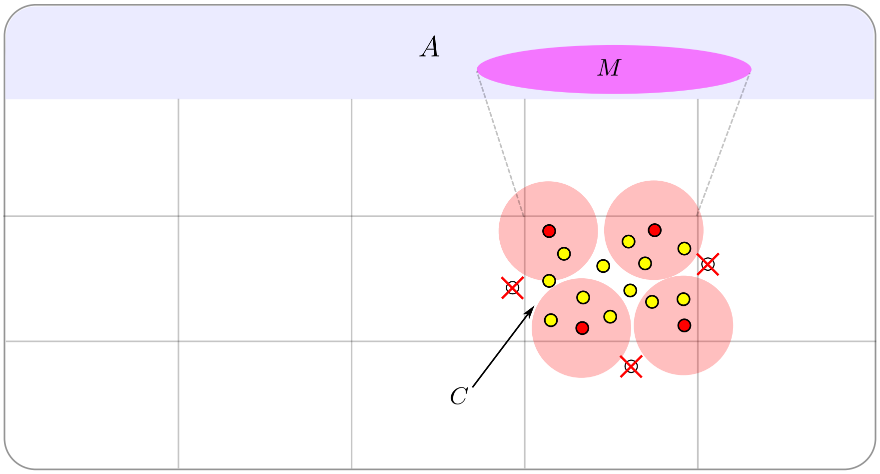

Figure 5: The structure exposed in the proof of Lemma 24. We have a set A of size $\mathcal{O}(k)$ that is an r-dominating set in G and $3r$-projections onto it are of constant size. We partition vertices of $V(G) - A$ according to $3r$-projection profiles on A and, provided Z is large, we find an equivalence class C (yellow and red vertices) consisting of a large (super-constant) number of elements of $Z - A$ with exactly the same $3r$-projection M on A. To find an irrelevant dominatee among them, we apply uniform quasi-wideness to highlight a large subset $C' \subseteq C$ (red vertices) that becomes $2r$-independent in $G - S$, for some set S (crossed out) of bounded size. We will later filter C' according to r-distance profiles on S, but this is not depicted.

$c_{\mathrm{nei}} c_{\mathrm{prj}} c_{\mathrm{apx}} k$ equivalence classes, there is an equivalence class C that has more than c' elements. From now on we will focus on the equivalence class C and find the irrelevant dominatee there.

The intuition now is that we have a really large set C of elements of Z that need to be r-dominated by any dominator of Z, and all of them have the same $3r$-projection (profile) on A. Since A is an r-dominating set, this $3r$-projection is nonempty, while the Projection Closure Lemma ensures us that it is also of size bounded by c_{prj}. Hence there is a very cheap way to r-dominate all the vertices of C: just take all the vertices of this $3r$-projection. We now would like to show that every optimal dominator of Z needs to, roughly, use this way of r-dominating C, as other ways would be suboptimal. Moreover, exclusion of any vertex z from C still forces such behavior, so the optimal r-domination of $C - \{z\}$ automatically forces r-domination of z as well due to equality of projection profiles.

This intuition is not entirely accurate, as we need to dig a bit more for the irrelevant dominatee. More precisely, this is the moment when we apply uniform quasi-wideness, to find a large $2r$-independent subset of C. Since $\mathcal{C}$ has bounded expansion, it is also nowhere dense, so it is also uniformly quasi-wide. Therefore, invoking uniform quasi-wideness for radius $2r$, there exists a constant s and a function $N(m)$ such that provided $|C| > N(m)$, we may find in polynomial time sets $S \subseteq V(G)$ and $C' \subseteq C - S$ with $|S| \leqslant s$, $|C'| > m$, and C' being $2r$-independent in $G - S$. We shall apply this for $m = (c_{\mathrm{prj}} + 1)(r + 2)^s$, so we accordingly set $c' := N((c_{\mathrm{prj}} + 1)(r + 2)^s)$. More precisely, by putting c' as above we can find sets S and C' with asserted properties and $|C'| > (c_{\mathrm{prj}} + 1)(r + 2)^s$.

We finally classify the vertices of C' according to their r-distance profiles on S. Precisely, since $|S| \leqslant s$, the number of different r-distance profiles on S is bounded by $(r + 2)^s$. Since

$|C'| > (c_{\mathrm{prj}} + 1)(r + 2)^s$, there is a subset $C'' \subseteq C'$ of size at least $c_{\mathrm{prj}} + 2$ such that all elements of C'' have the same r-distance profiles on S.

We now claim that any vertex $z \in C'$ is an irrelevant dominatee: precisely, for any fixed $z \in C'$ it holds that $Z - \{z\}$ is still a strict core. For this, take any optimal dominator D of $Z - \{z\}$ in G. If D in addition r-dominates z, then D is also an optimal dominator of Z and, by the assumption that Z is a strict core, D is an r-dominating set of G. We are left with examining the case when D does not r-dominate z. In this case we will derive a contradiction with the optimality of D.

For every $z' \in C'' - \{z\}$, let us fix any vertex $d_{z'} \in D$ that r-dominates z' and any path $P_{z'}$ of length at most r connecting z' with $d_{z'}$.

Claim 1. *For each $z' \in C'' - \{z\}$ we have that $P_{z'}$ is disjoint from $A \cup S$.*

Proof. Suppose that $P_{z'}$ intersects $A \cup S$ and let u be the vertex of $V(P) \cap (A \cup S)$ that is the closest (on $P_{z'}$) to z'. If $u \in A$ then, by the equality of $3r$-projection profiles of z and z' on A, we infer that d also r-dominates z; a contradiction. On the other hand, if $u \in S$ then the same contradiction follows from the equality of r-distance profiles of z and z' on S. ⌟

As $P_{z'}$ is disjoint from S, we have that $d_{z'}$ belongs to the r-neighborhood of z' in $G - S$. However, $C'' \subseteq C'$ is $2r$-independent in $G - S$, so these r-neighborhoods are pairwise disjoint. We infer that vertices $d_{z'}$ are pairwise different for $z' \in C'' - \{z\}$, so in particular there is at least $|C''| - 1 \geqslant c_{\mathrm{prj}} + 1$ of them.

Let M be the (common) $3r$-projection of vertices of C onto A. Recall that by the Projection Closure Lemma we have $|M| \leqslant c_{\mathrm{prj}}$, and M is nonempty since A is an r-dominating set in G. Now construct D' from D by removing all vertices $d_{z'}$ for $z' \in C'' - \{z\}$ and adding all vertices of M. Thus we remove at least $c_{\mathrm{prj}} + 1$ vertices and add at most c_{prj}, yielding $|D'| < |D|$. To obtain now a contradiction it suffices to prove the following.

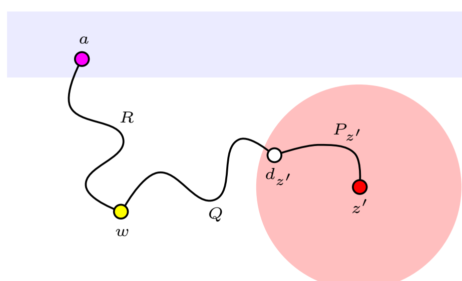

Figure 6: Situation in the proof of Claim 2.

Claim 2. *The set D' is a dominator of Z.*

Proof. Since M, the common $3r$-projection of vertices of C on A, is nonempty and has been explicitly included in D', we have that D' r-dominates every vertex of C, in particular z. Let us then take any $w \in Z - \{z\}$; we need to prove that D' still r-dominates w. Let d be a vertex of D that r-dominates w; such d exists since D was a dominator of $Z - \{z\}$ by assumption. If d is not among $\{d_{z'} \colon z' \in C'' - \{z\}\}$, then d is still contained in D' and then D' r-dominates w. So suppose otherwise: $d = d_{z'}$ for some $z' \in C'' - \{z\}$. Let Q be any path of length at most r from w to d and let R be any path of length at most r from w to a vertex $a \in A$; such path R exists since A

is an r-dominating set in G. Examine the walk W defined as the concatenation of paths $P_{z'}$, Q, and R, in this order. This is a walk of length at most $3r$ that connects z' with a vertex $a \in A$. Let a' be the first (closest to z') vertex of A on W. Then the prefix of W from z' to a' witnesses that $a' \in M$, so in particular a' is included in D'. Vertex a' cannot lie on $P_{z'}$, since $P_{z'}$ is disjoint from A by Claim 1. Therefore, a' lies on $Q \cup R$. However, all vertices of $Q \cup R$ are at distance at most r from w, because Q and R are paths of length at most r, hence $a' \in D'$ r-dominates w. ⌟

To summarize, Claim 2 implies that the case when D does not r-dominate z is impossible, implying that every optimal dominator of $Z - \{z\}$ is also an optimal dominator of Z, and hence it is an r-dominating set in G. For the argument to be applicable we needed to set

$$c' \coloneqq N((c_{\mathrm{prj}} + 1)(r + 2)^s) \qquad \text{and} \qquad c \coloneqq c_{\mathrm{nei}} c_{\mathrm{prj}} c_{\mathrm{apx}} c' + c_{\mathrm{prj}} c_{\mathrm{apx}},$$

where s and $N(\cdot)$ are margins governing uniform quasi-wideness of $\mathcal{C}$ for radius $2r$. We conclude that for such c Lemmas 24 and 25 hold.

6.3 Phase 2: reducing dominators

Once we we have reduced the number of dominatees to linear in k, we will reduce the number of dominators in one shot similarly as before. Let $Z \subseteq V(G)$ be a strict core of size at most $c_{\mathrm{core}}k$ obtained by applying the algorithm of Lemma 25; throughout this section we rename the constant c given by Lemma 25 to c_{core} to avoid confusion. The intuition is that now we care only about dominating vertices in Z. Hence, any two vertices of G that have the same r-neighborhood in Z are functionally equivalent, as they dominate the same subset of Z. Hence, we need to preserve only one of them.

To formalize this intuition, we introduce the following equivalence relation on $V(G)$: for $u, v \in V(G)$, we put

$$u \sim v \qquad \text{if and only if} \qquad N_r[u] \cap Z = N_r[v] \cap Z.$$

As we know from previous lectures, the number of equivalence classes of $\sim$ is bounded by $c_{\mathrm{nei}}|Z|$ for some constant c_{nei} depending only on $\mathcal{C}$ and r. This means that $\sim$ has at most $c_{\mathrm{nei}} c_{\mathrm{core}} \cdot k$ equivalence classes.

For each class C of $\sim$, arbitrarily select any its member $d_C \in C$. Let

$$W \coloneqq Z \cup \{d_C \colon C \text{ is an equivalence class of } \sim\}.$$

Finally, apply Shortest Paths Closure Lemma to W, yielding a set W' with $|W'| \leqslant c_{\mathrm{spc}}|W|$ for some constant c_{spc} depending only on $\mathcal{C}$ and r, such that $\mathrm{dist}_G(u, v) \leqslant r$ for $u, v \in W$ entails $\mathrm{dist}_{G[W']}(u, v) = \mathrm{dist}_G(u, v)$. Let $H = G[W']$; then $|W'| \leqslant c_{\mathrm{spc}} c_{\mathrm{nei}} c_{\mathrm{core}} \cdot k$, so we may set $c \coloneqq c_{\mathrm{spc}} c_{\mathrm{nei}} c_{\mathrm{core}}$. We are left with verifying that H and Z indeed satisfy all the required properties; the proof is essentially advanced symbol pushing.

Claim 3. *It holds that* $\mathrm{dom}_r(H, Z) = \mathrm{dom}_r(G)$.

Proof. We first prove that $\mathrm{dom}_r(H, Z) \geqslant \mathrm{dom}_r(G)$. Since H is an induced subgraph of G, every dominator of Z in H is also a dominator of Z in G, implying $\mathrm{dom}_r(H, Z) \geqslant \mathrm{dom}_r(G, Z)$. However, Z is a strict core in G, so $\mathrm{dom}_r(G, Z) = \mathrm{dom}_r(G)$ by Lemma 23.

We now prove that $\mathrm{dom}_r(H, Z) \leqslant \mathrm{dom}_r(G)$. Take any r-dominating set D in G, say of optimum size $\mathrm{dom}_r(G)$. By the construction of W, for every $d \in D$ there is some $d' \in W$ with $d \sim d'$. Let

$D' \coloneqq \{d' \colon d \in D\}$; clearly $|D'| \leqslant |D| = \mathrm{dom}_r(G)$. By the definition of $\sim$ we have that D' is still a dominator of Z in G, because every vertex $d' \in D'$ r-dominates the same vertices of Z as $d \in D$. Moreover, D' is a dominator of Z in H as well. This is because for every $d' \in D'$ and every $z \in Z$ that is r-dominated by d' in G, we have that $\mathrm{dist}_G(d', z) \leqslant r$ and $d', z \in W$, which implies that $\mathrm{dist}_H(d', z) = \mathrm{dist}_G(d', z) \leqslant r$ by the Shortest Paths Closure Lemma. We conclude that in H there is a dominator of Z of size at most $\mathrm{dom}_r(G)$, implying $\mathrm{dom}_r(H, Z) \leqslant \mathrm{dom}_r(G)$. ⌟

Claim 3 finishes the proof of Theorem 19.

6.4 Concluding remarks

The proof presented above combines many different tools that we have learnt to give a preprocessing algorithm for r-Dominating Set on any class of sparse graphs (formally, for any class of bounded expansion). The outcome of the algorithm is guarenteed to be of size linear in the target budget k, although the constant standing in front is, let's say, impractical.

It is natural to ask whether a similar preprocessing algorithm can be given for nowhere dense graph classes; the outcome, as usual, should be of size $f(\varepsilon) \cdot k^{1+\varepsilon}$ for any $\varepsilon > 0$, instead of linear in k. This is indeed true and the proof follows exactly the same strategy. The main difference is that all constants (apart from s, the bound on the number of removed vertices in uniform quasi-wideness) will be replaced by bounds of the form $f(\varepsilon) \cdot k^{\varepsilon}$ for any $\varepsilon > 0$. We need to be, however, careful that we do not use any reasonings that involve exponential functions of these "quasi-constants". In particular, we need to use neighborhood complexity for nowhere dense classes, which is significantly more difficult than for classes of bounded expansion, and when applying uniform quasi-wideness we need to know that the function $N(m)$ can be chosen to be polynomial in m — which does not follow from the proof we gave. During the next lectures we will give an introduction to tools using which these gaps can be filled.