VC-density and abstract cell decomposition for edge relation in graphs of bounded twin-width

Wojciech Przybyszewski
University of Warsaw, Warsaw, Poland

Abstract
We study set systems formed by neighborhoods in graphs of bounded twin-width. In particular, we prove that such classes of graphs admit linear neighborhood complexity, in analogy to previous results concerning classes with bounded expansion and classes of bounded clique-width. Additionally, we show how, for a given graph from a class of graphs of bounded twin-width, to efficiently encode the neighborhood of a vertex in a given set of vertices $A$ of the graph. For the encoding we use only a constant number of vertices from $A$. The obtained encoding can be decoded using FO formulas. This proves that the edge relation in graphs of bounded twin-width, seen as first-order structures, admits a definable distal cell decomposition. From this fact we derive that we can apply to such classes combinatorial tools based on the Distal cutting lemma and the Distal regularity lemma.

2012 ACM Subject Classification Mathematics of Computing → Discrete Mathematics → Graph Theory; Theory of computation → Logic → Finite Model Theory

Keywords and phrases graph theory, Vapnik-Chervonenkis density, twin-width, neighborhood complexity, abstract cell decomposition, regularity lemma

Funding This work is a part of project BOBR that has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 948057).

Acknowledgements I am deeply grateful to Dr. Szymon Toruńczyk, my master’s thesis supervisor, for introducing me to this area of research as well as for his assistance and dedicated involvement in every step throughout the process of preparing this paper.

1 Introduction

Twin-width is a graph width parameter recently introduced by Bonnet et al. in [6] with remarkable properties. In particular, it generalizes some of the previously examined graph classes (e.g. planar graphs, graphs excluding a fixed minor, graphs of bounded clique-width), while admitting good structural (e.g. being closed under FO-transductions), algorithmical (e.g. FO model checking can be done in linear time on graphs of bounded twin-width given a contraction sequence) and combinatorial (e.g. $\chi$-boundedness [4]) properties.

Our focus is on the combinatorial and logical complexity of set systems defined by neighborhoods in graphs of bounded twin-width. This continues a line of research studying other graph classes in this context. Namely, graphs of bounded clique-width admit linear neighborhood complexity, as it was proven in [20]. That means that for every graph $G$ of clique-width at most $c$ and any non-empty subset $A$ of its vertices there are at most $n_c|A|$ different neighborhoods of vertices of $G$ in $A$ (where the constant $n_c$ depends only on $c$). The same is also true for planar graphs and graphs excluding a fixed minor. More generally, the same holds for classes with bounded expansion [21], although they are not generalized by graphs of bounded twin-width. Our first main result states that graphs of bounded twin-width also admit linear neighborhood complexity.

Theorem 3.1. For every integer $t$, there is some $n_t$ such that for every graph $G$ of twin-width at most $t$ and every non-empty $A \subseteq V(G)$ we have $|N_G(A)| \leq n_t|A|$.
Our proof relies on properties of matrices of graphs of bounded twin-width shown in [6], as well as the celebrated Stanley-Wilf conjecture/Marcus-Tardos theorem [16]. We remark that Theorem 3.1 was proven independently by Bonnet et al. in [5].

The most well-known measure of complexity of set systems is VC-dimension, introduced in [27]. It is applied in a wide range of fields, including statistical learning theory and computational geometry. VC-density is a refinement of VC-dimension. Whereas a set system has bounded VC-dimension if and only if it has bounded VC-density, in general, VC-density is at most as large as the VC-dimension, and it turns out that the VC-density, not the VC-dimension, is the decisive measure for the combinatorial complexity of a family of sets [1]. For example, the VC-density of $\mathcal{S}$ governs the size of packings in $\mathcal{S}$ with respect to the Hamming metric [12] and is intimately related to the notions of entropic dimension [2] and discrepancy [18]. VC-density has also a number of applications in algorithmics, for example CONNECTED $k$-VERTEX COVER admits a kernel with $O(k^{1.5})$ vertices on classes of graphs with VC-density at most 1 [5]. Theorem 3.1 in particular implies that the VC-density of graphs of bounded twin-width is equal to 1.

In our second main result, we refine our analysis of set systems defined by neighborhoods in graphs of bounded twin-width, and turn our attention to their logical complexity. Namely, we investigate how to efficiently encode neighborhoods of vertices in graphs of bounded twin-width. The main result of our paper, in an informal formulation, states the following.

|$\blacktriangleleft$**Theorem (Informal version of Theorem 4.1).** For every integer $t$ there is a finite set $\Delta_t$ of first-order formulas such that for every ordered graph $(G, \preceq)$ of twin-width at most $t$ and for every $A \subseteq V(G)$ we can partition $V(G)$ into $O(|A|)$ parts such that vertices in every part have the same neighborhood in $A$ and every part is first-order definable using one of the formulas in $\Delta_t$ and parameters from $A$.

|$\blacktriangleleft$**Theorem (Cutting lemma).** For every set $L$ of $n$ lines in the real plane and every $1 < r < n$ there exists a $\frac{1}{r}$-cutting for $L$ of size $O(r^2)$. That is, there is a subdivision of the plane into generalized triangles (i.e. intersections of three half-planes) $\Delta_1, \ldots, \Delta_t$ so that the interior of each $\Delta_i$ is intersected by at most $\frac{n}{r}$ lines in $L$, and we have $t \leq Cr^2$ for a certain constant $C$ independent of $n$ and $r$.

This result provides a method to analyze intersection patterns in families of lines, and it has many generalizations to higher dimensional sets and/or to families of sets of more complicated shape than lines, for example families of algebraic or semialgebraic curves of bounded complexity [8]. Our version of cutting lemma for classes of graphs of bounded twin-width is as follows.
Theorem (Informal version of Theorem 5.4). For any \( t \) there is a constant \( C \) depending only on \( t \) such that following holds. For any graph \( G \) of twin-width at most \( t \), any \( A \subseteq V(G) \) of size \( n \) and any real \( 1 \leq r \leq n \) we can partition the vertices of \( V(G) \) into at most \( C t^{O(t)} \) sets \( X_1, \ldots, X_l \) such that the vertices in every \( X_i \) have almost the same neighborhood in \( A \). More precisely, there are at most \( \frac{n}{r} \) vertices \( a \in A \) for which there are \( u, v \in X_i \) with \((u, a) \in E(G) \) and \((v, a) \not\in E(G) \).

Let us remark that this theorem is obvious for \( r = 1 \) (we just take \( X_1 = V(G) \)) and the version for \( r = n \) follows from Theorem 3.1.

Another consequence of Theorem 4.1 is a strong version of the regularity lemma for graphs of bounded twin-width. The original Szemerédi’s regularity lemma is a fundamental result in graph combinatorics with many versions and applications in extremal combinatorics, number theory and computer science (see e.g. [25, 15]).

We prove a much stronger version of the regularity lemma for classes of graphs of bounded twin-width.

Theorem (Informal version of Theorem 5.4). For every \( t \) there exists \( C = C(t) \) such that: for every \( \varepsilon > 0 \) and for every graph \( G = (V, E) \) of twin-width at most \( t \), there is a partition \( V = V_1 \cup \ldots \cup V_k \) into non-empty sets, and a set \( \Sigma \subseteq [k] \times [k] \) with the following properties.

1. Polynomially bounded size of the partition: \( k = O(\varepsilon^{-\varepsilon}) \).
2. Few exceptions: \( |\bigcup_{(i,j) \in \Sigma} V_i \times V_j| \geq (1-\varepsilon)|V|^2 \).
3. \( \varepsilon \)-regularity: for all \( (i, j) \in \Sigma \) and all \( A \subseteq V_i, B \subseteq V_j \) with \( |A| \geq \varepsilon|V_i|, |B| \geq \varepsilon|V_j| \), one has
   \[
   |d(A, B) - d(V_i, V_j)| \leq \varepsilon,
   \]
   where \( d(X, Y) = \frac{|E(X,Y)|}{|X||Y|} \) and \( E(X, Y) \) is the set of edges with one endpoint in \( X \) and the other in \( Y \).

In general the bound on the size of the partition \( K \) is known to grow as an exponential tower of height \( \frac{t}{2} \). Recently several improved regularity lemmas were obtained in the context of definable sets in certain structures or in restricted families of structures (see e.g. [25, 15]).

We prove a much stronger version of the regularity lemma for classes of graphs of bounded twin-width.

2 Preliminaries

We denote by \([i]\) the set of integers \( \{1, \ldots, i\} \). If \( \mathcal{X} \) is a set of sets, we denote by \( \bigcup \mathcal{X} \) the union of them.
2.1 Graph definitions and notations

In this paper we investigate only undirected simple graphs, i.e. graphs with no multiple-edges nor self-loops. We denote by $V(G)$ the set of vertices of a given graph $G$ and by $E(G)$ the set of its edges.

For a graph $G = (V, E)$ and $A \subseteq V$ we denote by $G - A$ the graph $\{V \setminus A, \{(u, v) \in E : u, v \notin A\}\}$, i.e. the graph obtained from $G$ by removing all the vertices in $A$ and edges incident to them.

For $A \subseteq V(G)$ and $v \in V(G)$ we denote the neighborhood of $v$ in $A$ by $N^G_A(v)$, i.e. $N^G_A(v) = \{u \in A : (u, v) \in E(G)\}$. We omit a superscript $G$ whenever the graph is implicit from the context. We write $N(v)$ for $N_{V(G)}(v)$. Similarly, we denote the set of all neighborhoods in $A \subseteq V(G)$ by $N_G(A)$, i.e. $N_G(A) = \{N_A(u) : u \in V(G)\}$. We also write $\hat{N}_G(A)$ for the set of all non-empty neighborhoods in $A$, i.e. $\hat{N}_G(A) = N_G(A) \setminus \{\emptyset\}$.

For a graph $G = (V, E)$ and two subsets of its vertices $A, B \subseteq V$ we say that $A$ and $B$ are homogeneous, if either for every $v \in A$ and $u \in B$ we have $(v, u) \in E$ or for every $v \in A$ and $u \in B$ we have $(v, u) \notin E$. In particular, if $A \cap B = \emptyset$ then $A$ and $B$ can be homogeneous only if for every $v \in A$ and $u \in B$ we have $(v, u) \notin E$.

2.2 Matrix definitions and notations

For a matrix $M$ consisting of $m$ rows and $n$ columns (an $m \times n$ matrix) we denote by $M[i][j]$ its entry in the $i$'th row and $j$'th column (assuming that $1 \leq i \leq m$ and $1 \leq j \leq n$). A $p \times q$ submatrix $N$ of an $m \times n$ matrix $M$ (with $p \leq m$ and $q \leq n$) is any matrix formed by taking $p$ consecutive rows and $q$ consecutive columns of $M$. For an $m \times n$ matrix $M$ we denote by $M[i : j][k : l]$ (with $1 \leq i \leq j \leq m$ and $1 \leq k \leq l \leq n$) the submatrix of $M$ formed by rows from $i$ to $j$ (inclusive) and columns from $k$ to $l$ (inclusive).

An $m \times n$ matrix $M$ is vertical if every two rows of $M$ are equal, i.e. for any $1 \leq i, j \leq m$ we have $M[i][1 : n] = M[j][1 : n]$. Similarly, $M$ is horizontal if every two columns of $M$ are equal, i.e. for any $1 \leq i, j \leq n$ we have $M[1 : m][i] = M[1 : m][j]$. Observe that if a matrix is both vertical and horizontal, then it is constant. We say that a matrix is mixed if it is neither vertical nor horizontal. A corner is a $2 \times 2$ mixed matrix. In [6] it was proven that a matrix is mixed if and only if it contains a corner as a submatrix.

Given an $m \times n$ matrix $M$, a row-partition (resp. column-partition) is a partition of the rows (resp. columns) of $M$. Similarly, a row-division (resp. column-division) is a row-partition (resp. column-partition), where every part consists of consecutive rows (resp. columns). A $(k, l)$-division (or simply division) of a matrix $M$ is a pair $(R, C)$ of a row-division and a column-division with respectively $k$ and $l$ parts. A zone of a division $(R, C) = (\{R_1, \ldots, R_k\}, \{C_1, \ldots, C_l\})$ is any submatrix $R_i \cap C_j$ for $1 \leq i \leq k$ and $1 \leq j \leq l$.

A 0,1-matrix is a matrix with all its entries equal to 0 or 1. Given a 0,1-matrix $M$, a $t$-grid minor in $M$ is a $(t, t)$-division of $M$ in which every zone contains a 1. Given a 0,1-matrix $M$, a $t$-mixed minor in $M$ is a $(t, t)$-division of $M$ in which every zone is a mixed submatrix of $M$. A matrix is $t$-grid free (resp. $t$-mixed free) if it doesn’t contain a $t$-grid minor (resp. $t$-mixed minor).

If $G$ is an $n$-vertex graph and $\sigma$ is a total ordering of $V(G)$, say, $v_1, \ldots, v_n$, then $M_\sigma(G)$ denotes the adjacency matrix of $G$ in the order $\sigma$. Thus $M_\sigma(G)[i][j] = 1$ if $(v_i, v_j) \in E(G)$ and 0 otherwise. An ordered graph is a pair $(G, \preceq)$ such that $G$ is a graph and $\preceq$ is a total order on $V(G)$. We say that an ordered graph $(G, \preceq)$ is $t$-mixed free if $M_\preceq(G)$ is $t$-mixed free. A first-order formula of ordered graphs is a standard first-order formula of graphs that
can use additional ≤ symbol, which is interpreted as an order of an ordered graph.

2.3 Definition of twin-width of graphs

The twin-width of a given graph was first defined in [6]. That definition uses the notion of a trigraph, i.e. a triple \( G = (V, E, R) \) where \( E \) and \( R \) are two disjoint sets of edges on \( V \): the (usual) edges and the red edges. A trigraph \( (V, E, R) \) such that \( (V, R) \) has maximum degree at most \( d \) is a \( d \)-trigraph. Any graph \( (V, E) \) may be interpreted as the trigraph \( (V, E, \emptyset) \).

Given a trigraph \( G = (V, E, R) \) and two vertices \( u, v \) in \( V \), we define the trigraph \( G/u, v = (V', E', R') \) obtained by identifying \( u, v \) into a new vertex \( w \) as the trigraph on a new vertex-set \( V' = (V \setminus \{u, v\}) \cup \{w\} \) such that \( G - \{u, v\} = (G/u, v) - \{w\} \) and the following edges incident to \( w \):

- \((w, x) \in E'\) if and only if \((u, x) \in E\) and \((v, x) \in E\),
- \((w, x) \notin E' \cup R'\) if and only if \((u, x) \notin E\) and \((v, x) \notin E\),
- \((w, x) \in R'\) otherwise.

We say that \( G/u, v \) is a contraction of \( G \). If both \( G \) and \( G/u, v \) are \( d \)-trigraphs, \( G/u, v \) is a \( d \)-contraction.

A (tri)graph \( G \) is \( d \)-collapsible if there exists a sequence of \( d \)-contractions which contracts \( G \) to a single vertex. The minimum \( d \) for which \( G \) is \( d \)-collapsible is the twin-width of \( G \), denoted \( \text{tww}(G) \).

2.4 Theorems related to twin-width of graphs

The grid minor theorem for twin-width proven in [6] states, that there is a connection between twin-width of a graph and mixed minors of its matrix. One of its possible formulations is as follows:

> **Theorem 2.1.** If \( G \) is a graph of twin-width less than \( t \), then there is a total ordering on its vertices \( \sigma \) such that \( M_\sigma(G) \) is \( 2t + 2 \)-mixed free. On the other hand, if \( G \) is a graph and \( \sigma \) is a total ordering on its vertices such that \( M_\sigma(G) \) is \( k \)-mixed free, then \( \text{tww}(G) = 2^{2\Theta(k)} \).

A crucial tool for this proof is the Marcus-Tardos theorem about grid minors.

> **Theorem 2.2** ([16]). For every integer \( t \), there is some \( c_t \) such that every \( m \times n \) 0,1-matrix \( M \) with at least \( c_t \max(m, n) \) entries 1 has a \( t \)-grid minor.

Marcus and Tardos established Theorem 2.2 with \( c_2 = 2t^4 \binom{t}{2} \). Then Cibulka and Kynčl [11] decreased \( c_2 \) to \( 8/3(t + 1)^22^{4t} \). In the rest of this paper we will keep the notation \( c_t \) for these constants.

2.5 Vapnik-Chervonenkis density of graphs

In this section we introduce notions of Vapnik-Chervonenkis dimension (VC-dimension) and Vapnik-Chervonenkis density (VC-density) in the context of graphs. It is worth noting that it was first introduced in [27] and originated in statistics.

We say that \( A \subseteq V(G) \) is shattered if \( N_G(A) = 2^A \), i.e. for every subset of \( A \) there is a vertex \( v \in V(G) \) such that \( N_A(v) \) is precisely that subset of \( A \). For a graph \( G \) we define its VC-dimension as the size of the largest subset of \( V(G) \), which is shattered, and denote it by \( \dim(G) \). We denote the shatter function of a graph \( G \) by \( \pi_G : \mathbb{N} \to \mathbb{N} \) and define it as follows:

\[
\pi_G(k) = \max_{A \subseteq V(G), |A| \leq k} |N_G(A)|.
\]
For a non-empty class of graphs $\mathcal{C}$ we define its VC-dimension as the supremum over VC-dimensions of graphs in $\mathcal{C}$ and denote it by $\dim(\mathcal{C})$. Similarly, we denote its shatter function by $\pi_{\mathcal{C}}: \mathbb{N} \to \mathbb{N}$ and define it by:

$$\pi_{\mathcal{C}}(k) = \max\{\pi_G(k) : G \in \mathcal{C}\}.$$ 

Observe that this definition is correct, as $\pi_G(k) \leq 2^k$ for any graph $G$ and $k \in \mathbb{N}$.

We can bound the shatter function of a graph class $\mathcal{C}$ in terms of $\dim(\mathcal{C})$. This is stated in the shatter function lemma proven independently in [22] and [23].

▶ Lemma 2.3 (Shatter function lemma). If $\mathcal{C}$ is a graph class such that $\dim(\mathcal{C}) \leq d$, then $\pi_{\mathcal{C}}(k) = O(k^d)$.

For a class of graphs $\mathcal{C}$ we also define its VC-density (denoted by $vc(\mathcal{C})$) as follows:

$$vc(\mathcal{C}) = \begin{cases} \inf \{r \in \mathbb{R} : r > 0, \pi_{\mathcal{C}}(k) = O(k^r)\} & \text{if } \dim(\mathcal{C}) < +\infty, \\ +\infty & \text{otherwise.} \end{cases}$$

### 3 Neighborhood complexity of classes of graphs of bounded twin-width

In the following section we will prove that classes of graphs of bounded twin-width admit linear neighborhood complexity. This is formalized in Theorem 3.1.

▶ Theorem 3.1. For every integer $t$, there is some $n_t$ such that for every graph $G$ of twin-width at most $t$ and every non-empty $A \subseteq V(G)$ we have $|N_G(A)| \leq n_t|A|$.

The following definition and two lemmas will be useful for the proof.

▶ Definition 3.2. Let $A$ be an $m \times n$ 0,1-matrix, where $m, n \geq 2$. We define the corner matrix $B$ of matrix $A$ as the $(m-1) \times (n-1)$ matrix given by:

$$B[i][j] = \begin{cases} 1 & \text{if } M[i : (i+1)][j : (j+1)] \text{ is a corner,} \\ 0 & \text{otherwise.} \end{cases}$$

▶ Lemma 3.3. Let $A$ be a 0,1-matrix, which is $t$-mixed free. Then its corner matrix $B$ is $2t$-grid free.

Proof. Let us assume by contradiction, that $B$ admits a $(2t, 2t)$-division $(\mathcal{R}, \mathcal{C}) = (\{R_1, \ldots, R_{2t}\}, \{C_1, \ldots, C_{2t}\})$, which is a $2t$-grid minor. Every part $R_k$, which consists of $l$ rows of $B$, corresponds to $l + 1$ consecutive rows of $A$. Precisely, if $R_k$ consists of rows $i$ to $i + l - 1$ of $B$, then

![Figure 1 Example of a corner matrix](image-url)
entries in these rows depend only on rows \(i\) to \(i + l\) of \(A\). Moreover, whenever \(|k - m| > 1\), \(R_k\) and \(R_m\) correspond to disjoint sets of rows of \(A\) – in particular \(R_1, R_3, \ldots, R_{2t-1}\) correspond to \(t\) disjoint sets of rows of \(A\). The same happens for \(C_1, C_3, \ldots, C_{2t-1}\). These sets of rows and columns induce a \(t\)-mixed minor of \(A\). This is a contradiction to \(A\) being \(t\)-mixed free.

\[\therefore\]

\textbf{Lemma 3.4.} Let \(P\) be an \(m \times n\) 0,1-matrix such that all corners in \(P\) appear only in \(p\) different pairs of consecutive rows, i.e. there are \(p\) distinct integers \(1 \leq i_1, \ldots, i_p < m\) such that every corner in \(P\) is of a form \(P[i_j : (i_j + 1)][k : (k + 1)]\) for some \(1 \leq j \leq p\) and \(1 \leq k < n\). Then \(P\) has at most \(2^{p+1}\) different columns.

\[\therefore\]

\textbf{Proof.} We proceed by induction on \(p\). In the base case, i.e. when \(p = 0\), \(P\) does not contain any corner, so it is either horizontal or vertical. If \(P\) is horizontal, then all its columns are equal by definition. On the other hand, when \(P\) is vertical, then every column of \(P\) contains only 0s or only 1s. Therefore, \(P\) contains at most 2 different columns.

Now let us assume \(p \geq 1\) and the statement is true for \(p - 1\). Without loss of generality \(1 \leq i_1 < \ldots < i_p < m\). We can split \(P\) into two submatrices \(P_1 = P[1 : i_p][1 : n]\) and \(P_2 = P[(i_p + 1) : m][1 : n]\). All the corners of \(P_1\) are contained in \(p - 1\) different pairs of consecutive rows, so it contains at most \(2^p\) different columns. Moreover, \(P_2\) does not contain any corner at all, so it contains at most 2 different columns. Every column of \(P\) is made up using one column of \(P_1\) and one column of \(P_2\), so there are at most \(2^p \cdot 2 = 2^{p+1}\) different columns in \(P\).

Now we prove that graphs of bounded twin-width have linear neighborhood complexity.

\textbf{Proof of Theorem 3.1.} Let us take any graph \(G\) such that \(\text{tww}(G) \leq t\) and a non-empty subset \(A \subseteq V(G)\) of vertices of \(G\). By Theorem 2.1 there is a total ordering \(\sigma\) on \(V(G)\) such that \(M_\sigma(G)\) is \(2t + 2\)-mixed free. For the rest of this proof we will write \(M\) for \(M_\sigma(G)\).
Let $B \subseteq V(G)$ be a minimal subset of vertices of $G$ such that they represent all possible neighborhoods in $A$ apart from $\emptyset$. Formally, $B$ is a minimal subset of $V(G)$ such that $\overrightarrow{N}_G(A) = \{N_A(b) : b \in B\}$. Obviously $|N_G(A)| \leq |B| + 1$ so we need to bound the size of $|B|$. Let us also remark, that for any distinct $b, b' \in B$ we have $N_A(b) \neq N_A(b')$, as in the opposite case we could remove $b'$ from $B$ obtaining a smaller set with the property we want. Similarly, for any $b \in B$ we have $N_A(b) \neq \emptyset$.

Let us denote by $N$ the matrix that is obtained from $M$ by removing rows that correspond to vertices in $V(G) \setminus A$ and columns that correspond to vertices in $V(G) \setminus B$. Obviously, $N$ is an $|A| \times |B|$ $2t + 2$-mixed free matrix, as we obtain it by deleting some rows and columns of $M$, which is $2t + 2$-mixed free itself. We can see that any two consecutive columns of $N$ form a mixed matrix. Indeed, they cannot be equal, so the matrix formed by them is not vertical. It is not horizontal either, as there is no column in $N$ containing only zeros. Therefore, for any two consecutive columns of $N$ there is a corner contained in them. Let us denote by $C$ the corner matrix of $N$ (we assume $|A|, |B| \geq 2$ as the opposite case is trivial).

From what we observed so far, we know that in every column of $C$ there is at least one entry with 1.

By Lemma 3.4 we obtain that every set of $2^{4t+4}$ consecutive columns of $C$ has at least $c_{4t+4}$ non zero rows. This is because every such set corresponds to a set of $2^{4t+4} + 1$ consecutive columns in $N$ (and columns of $N$ are pairwise different).

Suppose $C$ has at least $2^{4t+4}(|A| - 1)$ columns. Pick any partition of its columns into $|A| - 1$ disjoint sets $(B_1, \ldots, B_{|A|-1})$, of at least $2^{4t+4}$ consecutive columns each. Consider the $((|A| - 1) \times (|A| - 1))$ matrix $C'$ given by:

$$C'[i][j] = \begin{cases} 1 & \text{if there is a column in } B_j \text{ that has entry 1 in } i\text{'th row,} \\ 0 & \text{otherwise.} \end{cases}$$

The matrix $C'$ has at least $c_{4t+4}$ entries with 1 in every column. Therefore, by Theorem 2.2, it admits a $4t + 4$-grid minor, which trivially induces a $4t + 4$-grid minor of $C$. This is a contradiction with Lemma 3.3, which states that $C$ is a $4t + 4$-grid free. Therefore $C$ has at most $2^{4t+4}(|A| - 1) - 1$ columns, so $G$ has at most $2^{4t+4}(|A| - 1) + 1$ neighborhoods in $|A|$. Bearing in mind the separate case for $|A| = 1$, we can set $n_e = 2^{4t+4} = 2^{2O(t)}$.

\textbf{Corollary 3.5.} Let $\mathcal{C}_t$ be a class of all graphs of twin-width at most $t$. Then $\text{vc}(\mathcal{C}_t) = 1$.

**Proof.** Theorem 3.1 trivially implies $\text{vc}(\mathcal{C}_t) \leq 1$, so it suffices to show $\text{vc}(\mathcal{C}_t) \geq 1$. By $G_k$ we denote the graph on vertices $[2k]$ with edges $\{(2i - 1, 2i) : i \in [k]\}$, i.e. $G_k$ is a sum of $k$ non-incident edges. Let us also take $A_k = \{2i : i \in [k]\}$. Obviously $G_k$ has twin-width 0, so $G_k \in \mathcal{C}_t$ for any $t \geq 0$. On the other hand, $|N_{G_k}(A_k)| = |A_k| + 1$, which implies $\text{vc}(\mathcal{C}_t) \geq 1$.

\section{Efficient encoding of neighborhoods in classes of graphs of bounded twin-width}

In Section 3 we have proven that for any class $\mathcal{C}$ of graphs of bounded twin-width, any $G \in \mathcal{C}$, and any subset $A \subseteq V(G)$, vertices of $G$ induce at most $O(|A|)$ different neighborhoods in $A$. In the proof we used Theorem 2.1 – namely we assumed that there is an order $\preceq$ on $V(G)$ such that $M_z(G)$ is $2t + 2$-mixed free. In this section we discuss how to efficiently encode the neighborhood in $A$ of a given vertex $v \in V(G)$. Intuitively, we want to partition $V(G)$ into at most $m|A|$ parts, such that any two vertices $v, v'$ in the same part $P$ satisfy $N_A(v) = N_A(v')$. Moreover, we want to assign to every part $P$ a sequence $\overline{v} \in A^k$ and a
Vertices in every blue cell have the same neighborhood in $A$ and every such cell is definable by some formula $\varphi \in \Delta$ with parameters from $A$, for some fixed finite set of formulas $\Delta$.

formula $\varphi$ from a finite set $\Delta$ of first-order formulas in the language of ordered graphs with $k+1$ parameters, such that $P = \{v \in V(G) : \varphi(v, \pi)\}$. Then, the neighborhood of every $v \in V(G)$ can be encoded as the sequence $\pi$ and a formula $\varphi$ assigned to its part. Obviously, the constants $k$, $m$ and the set $\Delta$ should depend only on the class $C$. This intuition is stated formally in Theorem 4.1 and depicted in Figure 3.

▶ Theorem 4.1. For every integer $t$ there are constants $k_t, m_t$ and a finite set of first-order formulas in the language of ordered graphs with $k_t + 1$ parameters $\Delta_t$, such that for any $t$-mixed free ordered graph $(G, \preceq)$ and any non-empty $A \subseteq V(G)$ there are functions $f : V(G) \to A^{k_t}, g : V(G) \to \Delta_t$ such that:

- $|f(V(G))| \leq m_t|A|$;
- for all $v \in V(G)$ it holds $(G, \preceq) \models g(v)(v, f(v))$;
- for all $v, v' \in V(G)$ if $(G, \preceq) \models g(v)(v', f(v))$ then $N_A(v) = N_A(v')$.

We start by introducing a crucial definition for this part.

▶ Definition 4.2. Let $M$ be an $m \times n 0 \rightarrow 1$ matrix. We define the corner-profile of column $j$ (with $1 \leq j < n$) as the subset $\text{Prof}(j) \subseteq [m-1]$ such that for any $t$-mixed free ordered graph $(G, \preceq)$ and any non-empty $A \subseteq V(G)$ there are functions $f : V(G) \to A^{k_t}, g : V(G) \to \Delta_t$ such that:

- $|f(V(G))| \leq m_t|A|$;
- for all $v \in V(G)$ it holds $(G, \preceq) \models g(v)(v, f(v))$;
- for all $v, v' \in V(G)$ if $(G, \preceq) \models g(v)(v', f(v))$ then $N_A(v) = N_A(v')$.

We will use additional notation for an ordered graph $(G, \preceq)$. A vertex $v'$ is the successor of vertex $v$ in $\preceq$ (denoted by $s(v)$) if $v \neq v'$, $v' \preceq v$, and for all $u \in V(G)$ we have either $u \preceq v$ or $v' \preceq u$. Two vertices $v, v' \in V(G)$ are adjacent in $\preceq$ if $v' = s(v)$ or $v = s(v')$. For a vertex $v$ we also denote by $S^k(v)$ the set of $k$ consecutive vertices after $v$ together with $v$, i.e. $S^k(v) = \{s^l(v) : 0 \leq l \leq k\}$.
Proposition 4.3. Let \((G, \preceq)\) be a \(t\)-mixed free ordered graph and let \(v \in V(G)\) satisfy \(|\Prof(v)| \geq 2t\). There are a constant \(f_t\) depending only on \(t\), a sequence of vertices \((a_1, \ldots, a_t) \in \Prof(v)^t\), and a first-order formula in the language of ordered graphs \(\varphi(x, y_1, \ldots, y_t)\) such that the length of \(\varphi\) is bounded by \(f_t\) and \((G, \preceq) \models \varphi(x, a_1, \ldots, a_t) \iff x = v\).

Proof. As \(|\Prof(v)| \geq 2t\) we can pick a subset \(P \subseteq \Prof(v)\) of size \(t\) such that every two vertices \(w, w' \in P\) are not adjacent in \(\preceq\). Let us denote \(S = \{u \in V(G) : P \subseteq \Prof(u)\}\). Obviously \(v \in S\) and we can easily see that \(|S| < 2t\). Indeed, if \(|S| \geq 2t\), then we could pick \(S' \subseteq S\) of size \(t\) such that every two vertices in \(S'\) are not adjacent in \(\preceq\). Then pairs of rows \(\{(w, s(w)) : w \in P\}\) and pairs of columns \(\{(u, s(u)) : u \in S'\}\) would induce a \(t\)-mixed minor of \(M_\preceq(G)\). That is a contradiction with the assumption that \((G, \preceq)\) is \(t\)-mixed free. Therefore, we take \(\varphi\) which just says that \(v\) is the \(k\)th vertex in order \(\preceq\) such that \(P \subseteq \Prof(v)\). This can be done within the constraints in the statement, as we can use vertices in \(P\) as \(a_1, \ldots, a_t\) and we can define the successor of a given vertex using \(\preceq\). Moreover, the size of \(\varphi\) depends only on \(t\) (as we discussed that \(k\) is smaller than \(2t\)), so it can be bounded by some constant \(f_t\).

The assumption that one vertex has a large corner-profile is quite strong, in the sense that we can easily imagine a matrix which admits a large mixed minor, but its every column has a small corner-profile (e.g. of a constant size). However, Proposition 4.3 can be leveraged to the case when a set of consecutive vertices has a large corner-profile, which is formalized in Proposition 4.4.

Proposition 4.4. Let \((G, \preceq)\) be a \(t\)-mixed free ordered graph and let \(v \in V(G)\) satisfy \(|\Prof(S^l(v))| \geq 2t\) for some integer \(l\). There are a constant \(f_t\) depending only on \(t\), a sequence of vertices \((a_1, \ldots, a_t) \in \Prof(S^l(v))^t\), a first-order formula in the language of ordered graphs \(\varphi(x, y_1, \ldots, y_t)\), and a vertex \(v' \in S^l(v)\) such that the length of \(\varphi\) is bounded by \(f_t\) and \((G, \preceq) \models \varphi(x, a_1, \ldots, a_t) \iff x = v'\).

Proof. Similarly as in the proof of Proposition 4.3 we can pick a subset \(P \subseteq \Prof(S^l(v))\) of size \(t\) such that every two vertices \(w, w' \in P\) are not adjacent in \(\preceq\). Let \(v_0\) be the smallest vertex of \(G\) in \(\preceq\) and \(r_0 = \min\{n \in \mathbb{N} : P \subseteq \Prof(S^n(v_0))\}\), i.e. \(r_0\) is the smallest integer such that first \(r_0 + 1\) vertices of \(G\) in \(\preceq\) have \(P\) as a subset of their corner-profile. This is obviously well defined as \(P \subseteq S^l(v)\). Now we set \(v_1 = s^{r_0 + 1}(v_0)\) and \(r_1 = \min\{n \in \mathbb{N} : P \subseteq \Prof(S^n(v_1))\}\). We define \(v_2, v_3, \ldots\) analogically, as long as some \(v_j\) or \(r_j\) is not defined (i.e. \(s^{r_j - 1}(v_{j-1})\) is the maximal vertex of \(G\) in \(\preceq\) or...
Figure 5 Example of how \( v_i \) are set in the proof of Proposition 4.4.

\( \{ n \in \mathbb{N} : P \subseteq \text{Prof}(S^n(v_j)) \} \) is empty. We clearly have \( j < 2t \). Indeed, if \( j \geq 2t \) then the sets of columns \( S^{r_0+1}(v_0), S^{r_2+1}(v_2), \ldots, S^{r_{2t-2}+1}(v_{2t-2}) \) are pairwise disjoint. Moreover, together with pairs of rows \( \{(u, s(u)) : u \in P\} \) they induce a \( t \)-mixed minor of \( M_{\lessdot}(G) \), which is a contradiction.

If \( v_{j-1}, r_{j-1} \) is the last pair, which was well defined above, and we also managed to define \( v_j \) but not \( r_j \), then we can’t have \( v_j \lessdot v \). That is because we have \( P \subseteq \text{Prof}(S^i(v)) \), so in this case \( r_j \) would be well defined. Therefore, there is some \( p \) such that \( v \in S^{r_p}(v_p) \). We also have \( S^{r_p}(v_p) \in S^j(v) \). Indeed, we can’t have \( S^{r_p}(v_p) \lessdot v \) by the definition of \( p \) and if we had \( s^j(v) \lessdot S^{r_p}(v_p) \), then we would have \( P \subseteq S^j(v) \subseteq S^{r_p-1}(v_p) \), which contradicts the definition of \( r_p \). We pick the vertex \( S^{r_p}(v_p) \) as our \( v' \) from the statement. That is because it is either the maximal vertex of \( G \) (in which case it obvious that we can encode it even without parameters \( a_1, \ldots, a_t \)) or the predecessor of \( v_{p+1} \). However, \( v_{p+1} \) is definable by a first-order formula of ordered graphs with \( a_1, \ldots, a_t \) being the vertices of \( P \) in some order. This is because the definition of each \( v_i \) can be easily transformed into a first-order formula of ordered graphs, and the size of formula for \( v_p \) can be bounded by some constant \( f_t \) because \( p \leq 2t \). Finally, as we mentioned in the proof of Proposition 4.3, the successor relation is definable, so indeed we can define \( S^{r_p}(v_p) \) as the predecessor of \( v_{p+1} \).

Let us remark that Proposition 4.4 assures that if we have a set of consecutive vertices with a large corner-profile, then we can encode some vertex in it using a first-order formula. However, we cannot control which vertex from the given set is to be encoded. It turns out not to be that problematic, so we can start the proof of Theorem 4.1.

Proof of Theorem 4.1. Let us take a \( t \)-mixed free ordered graph \((G, \lessdot)\) and some non-empty \( A \subseteq V(G) \). Similarly as in the proof of Theorem 3.1 we can remove from \( M_{\lessdot}(G) \) rows
that correspond to vertices in \( V(G) \setminus A \), thus obtaining a \( t \)-mixed free \(|A| \times |V(G)|\) matrix \( M \).

We define \( v_0 \) to be the minimal vertex of \( G \) in \( \preceq \) and \( r_0 = \min \{ n \in \mathbb{N} : |\text{Prof}(S^n(v_0))| \geq c_{2t} \} \).

Next, we set \( v_1 = s_{r_0+1}(v_0) \) and \( r_1 = \min \{ n \in \mathbb{N} : |\text{Prof}(S^n(v_1))| \geq c_{2t} \} \). We proceed this way with \( v_2, r_2, v_3, \ldots \). If at some point we manage to define \( v_l \) but the set \( \{ n \in \mathbb{N} : |\text{Prof}(S^n(v_l))| \geq c_{2t} \} \) is empty, then we define \( r_l = \max \{ n \in \mathbb{N} : S^n(v_l) \in V(G) \} \), i.e. \( S^{r_l}(v_l) = \{ u \in V(G) : v_l \preceq u \} \). In this way we define \( l+1 \) pairs \( v_0, r_0, \ldots, v_l, r_l \). Using exactly the same argument as in the proof of Theorem 3.1 we obtain \( l + 1 \leq 2 \cdot 2^{2t} |A| = O(|A|) \).

As \( c_{2t} \geq 2t \), then in every \( S^{r_j}(v_j) \) there is some \( v_j' \), which is first-order definable. This is a bit more tricky than in the proof of Proposition 4.4, as now any two consecutive rows of \( M \) don’t necessarily correspond to vertices of \( V(G) \) which are adjacent in \( \preceq \). However, we are still able to define a corner by explicitly naming both its rows. Therefore, the formula \( \varphi_j \) which defines \( v_j' \) needs to use \( 2t \) parameters which correspond to \( t \) vertices from \( \text{Prof}(S^{r_j}(v_j)) \) and their successors in \( \preceq \). There is also a special case for \( S_{r_j}(v_j) \), as we can’t assume \( |\text{Prof}(S_{r_j}(v_j))| \geq 2t \). Nevertheless, the vertex \( s^{r_j}(v_j) \) is definable as the maximal vertex of \( G \), so we can set \( v_j' = s^{r_j}(v_j) \). As \( M \) is \( t \)-mixed free, then the size of each \( \varphi_j \) is bounded by some function of \( t \) only.

If for any \( S^{r_j}(v_j) \), \( \text{Prof}(S^{r_j}(v_j)) \geq 2t \), then we can encode \( S^{r_j}(v_j) \) and \( v_{j+1} \) (if it exists) as well. Again, we need to use it for formulas with \( 2t \) parameters of size bounded by some function of \( t \).

Now we will show how to define neighborhoods of remaining vertices. Every given vertex \( u \) satisfies \( w \prec u \prec w' \) for some \( w, w' \) already defined in previous steps. Moreover, there are at most \( 2c_{2t} \) corners between columns of \( w \) and \( w' \). By analyzing the proof of Lemma 3.4 we know that if some vertex \( w \prec u \prec w' \) has exactly the same neighborhood on corners between \( w \) and \( w' \), then their columns in \( M \) are equal, so they have the same neighborhood in whole \( A \). Therefore, we can define sets of vertices between \( w \) and \( w' \) in \( \preceq \), that have specific values on \( 2c_{2t} \) corners between \( w \) and \( w' \) using formulas with \( 2t + 2t + 4c_{2t} \) parameters of size bounded by some function of \( t \). Finally, we obtain the statement of Theorem 4.1 with constants \( k_t = 4c_{2t} + 4t, m_t = 2^{2t+1}, \) and \( \Delta_t \) consists of first-order formulas of ordered graphs with \( 4c_{2t} + 4t + 1 \) parameters, whose length can be bounded by some function of \( t \), which can be derived from this proof.

Constants \( k_t \) obtained in the proof of Theorem 4.1 are of order \( 2^{O(t)} \). If we allow partitioning \( V(G) \) into more than \( O(|A|) \) parts, we can obtain formulas with \( k_t = O(t) \). Moreover, by adding additionally a constant number of vertices from \( A \) we can obtain just one formula \( \psi \) instead of a finite set \( \Delta \).

**Proposition 4.5.** For every integer \( t \) there are a constant \( l_t = O(t) \) and a first-order formula in the language of ordered graphs with \( l_t + 1 \) parameters \( \psi_t \), such that for any \( t \)-mixed free ordered graph \( (G, \preceq) \) and any non-empty \( A \subseteq V(G) \) there is a function \( f : V(G) \rightarrow A^{l_t} \) such that:

= for all \( v \in V(G) \) it holds \( (G, \preceq) \models \psi_t(v, f(v)) \);

= for all \( v, v' \in V(G) \) if \( (G, \preceq) \models \psi_t(v', f(v)) \) then \( N_A(v) = N_A(v') \).

**Proof.** If \( |A| = 1 \) then the statement is obvious. Therefore, assume \( |A| \geq 2 \). We start by repeating the proof of Theorem 4.1, but for each \( v_j \) we define \( r_j = \min \{ n \in \mathbb{N} : |\text{Prof}(S^n(v_j))| \geq 2t \} \). In this way we may obtain more than \( O(|A|) \) pairs \( v_j, r_j \), but we get a finite set of formulas \( \Gamma_t \) with at most \( 12t \) parameters each, such that for every \( u \in S^{r_j}(v_j) \) we can encode neighborhood of \( u \) in \( A \) using one of the formulas in \( \Gamma_t \) (with suitable parameters from \( A \)). Moreover, the size of \( \Gamma_t \) does not depend on \( t \) (only formulas in \( \Gamma_t \) do), so by
using additionally \( \log |\Gamma_t| \) parameters from \( A \) we can write one formula \( \psi_t \), which uses these additional parameters to pick a correct formula from \( \Gamma_t \) (and here we need the assumption \( |A| \geq 2 \)).

5 Combinatorial tools for graphs of bounded twin-width

In this Section we derive powerful combinatorial consequences of Theorem 4.1 proved in the previous section. Namely, we observe that our theorem implies that ordered graphs of bounded twin-width allow *abstract cell decompositions* studied in model theory [9]. Combining with results from [9] and [10], this yields immediate powerful structural properties of graphs of bounded twin-width, such as the Distal cutting lemma and the Distal regularity lemma. We start this section with a few more definitions related to logic and model theory.

5.1 Structures and binary structures

In this paper we investigate only first-order structures over relational signatures.

Our notations are standard. We denote first-order structures by blackboard bold letters \( \mathbb{M}, \mathbb{N}, \) etc., and use letters \( M, N, \) etc. to denote their domains. We denote elements of domains and variables in formulas by small letters \( a, x, \) etc., and use small letters with a bar \( \bar{\pi}, \bar{x}, \) etc. to denote tuples of elements or variables. Then we use \( |\bar{x}| \) to denote the length of the tuple \( \bar{x} = (x_1, \ldots, x_n) \).

If \( \mathbb{M} \) is a structure and \( \varphi(\bar{x}; \bar{y}) \) is a formula in the language of \( \mathbb{M} \), then for \( \bar{v} \in M^{\bar{y}} \), by \( \varphi(M; \bar{v}) \) we denote the subset of \( M^{\bar{y}} \) defined by \( \varphi(\bar{v}; \bar{y}) \), namely \( \varphi(M; \bar{v}) = \{ b \in M^{\bar{y}} : \mathbb{M} \models \varphi(\bar{b}; \bar{y}) \} \).

For a signature \( \Sigma \) we say that \( \Sigma^+ \) is its *expansion* by unary predicates, if \( \Sigma^+ \) is the union of \( \Sigma \) and a set of unary relation symbols. For a \( \Sigma^+ \)-structure \( A^+ \) we say that a \( \Sigma \)-structure \( A \) is the *reduct* of \( A^+ \), if it is obtained from \( A^+ \) by forgetting relations in \( \Sigma^+ \setminus \Sigma \). We say that a \( \Sigma^+ \)-structure \( A^+ \) is a *monadic lift* of a \( \Sigma \)-structure \( A \), if \( A \) is the reduct of \( A^+ \).

We say that a \( \Sigma \)-structure \( A \) is a *binary structure*, if \( \Sigma \) is a signature consisting only of binary relational symbols.

5.2 Definition of twin-width of matrices and binary structures

The twin-width is also defined in [6] for matrices that have entries over a finite alphabet with a special additional value \( r \) (for red) representing errors made along the computation. The red number of a matrix is the maximum number of red entries taken over all rows and all columns. Given an \( m \times n \) matrix \( M \) and two columns \( C_i \) and \( C_j \), the contraction of \( C_i \) and \( C_j \) is obtained by deleting \( C_j \) and replacing every entry \( M[k][i] \) of \( C_i \) by \( r \) whenever \( M[k][i] \neq M[k][j] \). The same contraction operation is defined for rows.

A matrix \( M \) has twin-width at most \( k \) if one can perform a sequence of contractions starting with \( M \) and ending in some \( 1 \times 1 \) matrix in such a way that all matrices occurring in the process have red number at most \( k \). Note that when \( M \) has twin-width at most \( k \), one can reorder its rows and columns in such a way that every contraction will identify consecutive rows and columns. The reordered matrix is then called \( k \)-**twin-ordered**.

The symmetric twin-width of an \( n \times n \) matrix \( M \) is defined similarly, except that the contraction of rows \( i \) and \( j \) (resp. columns \( i \) and \( j \)) is immediately followed by the contraction of columns \( i \) and \( j \) (resp. rows \( i \) and \( j \)). Again, if \( M \) has symmetric twin-width at most \( k \), we can reorder its rows and columns in such a way that every contraction will identify consecutive rows and columns. The reordered matrix is then called \( k \)-**symmetric-twin-ordered**.
For a binary structure $S$ with a domain $S = \{v_1, \ldots, v_n\}$ over $h$ binary relations $E^1, \ldots, E^h$ we build its adjacency matrix $M$ in such a way that $M[i][j] = (e_1, \ldots, e_h) \in \{-1, 0, 1, 2\}^h$ where every $e_p$ is chosen accordingly. Namely:

- if $E^p(v_i, v_j)$ and $E^p(v_j, v_i)$ we set $e_p = 2$;
- if $E^p(v_i, v_j)$ and $\neg E^p(v_j, v_i)$ we set $e_p = 1$;
- if $\neg E^p(v_i, v_j)$ and $E^p(v_j, v_i)$ we set $e_p = -1$;
- if $\neg E^p(v_i, v_j)$ and $\neg E^p(v_j, v_i)$ we set $e_p = 0$.

The twin-width of a binary structure is the symmetric twin-width of the so-built adjacency matrix.

Let us assume that for some binary structure $S$ there is a total order $\sigma$ on $S$ such that the adjacency matrix of $S$ in this order is $k$-symmetric-twin-ordered. It is straightforward to see, that if we add $\sigma$ to $S$ as a new binary relation $E^{h+1}$ (thus obtaining a new structure $S^+$), then the adjacency matrix of $S^+$ in order $\sigma$ is still $k$-symmetric-twin-ordered.

Observe, that whenever a graph $G$ has twin-width at most $t$ in the sense of definition in Section 2.3 then its twin-width as a binary structure in the sense of this definition is at most $t + 2$. Let $\preceq$ be a total order on $V(G)$ such that the adjacency matrix of $G$ with rows and columns ordered by $\preceq$ is $(t + 2)$-symmetric-twin-ordered. We clearly see that the binary structure $(G, \preceq)$ with two binary relations $E$ and $\preceq$ also has twin-width at most $t + 2$. Moreover, by [6, Theorem 10] we obtain that $M_{\preceq}(G)$ is $2t + 6$-mixed free. For a graph $G$ of twin-width at most $t$ we call a total order $\preceq$ on its vertices, that satisfies properties above, a $t$-proper order for $G$.

### 5.3 NIP and monadically NIP structures

The definitions in this section are based on [10] and [19].

An important notion of NIP (standing for "not the independence property") models was introduced by Shelah in his work on the classification program [24]. As was observed early on in [14], the original definition of NIP is equivalent to the following one.

**Definition 5.1.** A structure $M$ is an NIP structure (or a dependent structure) if for every formula $\varphi(\overline{x}; \overline{y})$ with two tuples of parameters $\overline{x}$ and $\overline{y}$, there is some constant $k$, such that for all tuples $a_i \in M[\overline{x}](i \in [l])$, $b_I \in M[\overline{y}](I \subseteq [l])$ if

$$M \models \varphi(a_i, b_I) \iff i \in I,$$

then $l \leq k$.

A stronger notion of dependence arises when one allows to apply arbitrary monadic lifts to the structure before using the formula $\varphi$. This variant is called monadic NIP. The expressive power gained by the monadic lift is so strong that tuples of free variables can be replaced by single free variables in the above definition [3].

**Definition 5.2.** A $\Sigma$-structure $M$ is monadically NIP (or monadically dependent) if for every expansion $\Sigma^+$ of $\Sigma$ by unary predicate symbols and every first-order $\Sigma^+$-formula $\varphi(x, y)$ there exists an integer $k$ such that for every monadic lift $M_+^+$ of $M$ and for all elements $a_i \in M(i \in [l])$, $b_I \in M(I \subseteq [l])$ if

$$M \models \varphi(a_i, b_I) \iff i \in I,$$

then $l \leq k$. 

5.4 Definable sets and abstract cell decompositions

Let $\mathcal{M}$ be a fixed model. For sets $A, X \subseteq M^d$ we say that $A$ crosses $X$ if both $X \cap A$ and $X \cap (M \setminus A)$ are nonempty.

For a formula $\varphi(\bar{x}; \bar{y})$ and a set $S \subseteq M^{[\bar{x}]}$ we say that a subset $A \subseteq M^{[\bar{y}]}$ is $\varphi(\bar{x}; S)$ definable if $A = \varphi(M; \pi)$ for some $\pi \in S$. For a set $A \in M^{[\bar{y}]}$ we say that $\varphi(\pi, S)$ crosses $A$ if some $\varphi(\pi, S)$ definable set crosses $A$.

Given a finite set $S \subseteq M^{[\bar{x}]}$, a finite family $\mathcal{F}$ of subsets of $M^{[\bar{x}]}$ is called an abstract cell decomposition for $\varphi(\bar{x}, S)$ if $M^{[\bar{x}]} = \bigcup \mathcal{F}$ and every $\Delta \in \mathcal{F}$ is not crossed by $\varphi(\bar{x}; S)$. An abstract cell decomposition for $\varphi(\bar{x}; S)$ is an assignment $\mathcal{T}$ that to each finite set $S \subseteq M^{[\bar{x}]}$ assigns an abstract cell decomposition $\mathcal{T}(S)$ for $\varphi(\bar{x}; S)$.

Observe that every $\varphi(\bar{x}; S)$ admits an obvious abstract cell decomposition, with $\mathcal{T}(S)$ consisting of the atoms in the Boolean algebra generated by the $\varphi(\pi, S)$-definable sets. In general, defining these cells would require longer and longer formulas when $S$ grows, and we want to avoid this possibility. Therefore, we say that an abstract cell decomposition $\mathcal{T}$ for $\varphi(\bar{x}; S)$ is weakly definable if there is a finite set of formulas $\Psi(\bar{x}; \bar{y_1}, \ldots, \bar{y_k})$ with $|\bar{y}| = |\bar{y_1}| = \ldots = |\bar{y_k}|$ such that for any finite $S \subseteq M^{[\bar{x}]}$, every $\Delta \in \mathcal{T}(S)$ is $\psi(\pi; S^k)$-definable (i.e. $\Delta = \psi(M; \pi_1, \ldots, \pi_k)$ for some $\pi_1, \ldots, \pi_k \in S$ and $\psi \in \Psi$). In this case we say that $\Psi(\bar{x}; \bar{y_1}, \ldots, \bar{y_k})$ weakly defines $\mathcal{T}$.

For combinatorial applications it is useful to have control over sets appearing in $\mathcal{T}(S)$ in a definable way. Therefore, we say that an abstract cell decomposition $\mathcal{T}$ for $\varphi(\bar{x}; S)$ is definable if it is weakly definable by some $\Psi(\bar{x}; \bar{y_1}, \ldots, \bar{y_k})$ and if for every $S \subseteq M^{[\bar{x}]}$ and each $\Psi(\bar{x}; S^k)$-definable $\Delta \subseteq M^{[\bar{x}]}$ there is a set $\mathcal{I}(\Delta) \subseteq M^{[\bar{x}]}$, uniformly definable in $\Delta$, such that

$$\mathcal{T}(S) = \{ \Delta \in \Psi(\bar{x}; S) : \mathcal{I}(\Delta) \cap S = \emptyset \}. $$

By the uniform definability of $\mathcal{I}(\Delta)$ we mean that for every $\psi(\bar{x}; \bar{y_1}, \ldots, \bar{y_k}) \in \Psi(\bar{x}; \bar{y_1}, \ldots, \bar{y_k})$, there is a formula $\theta_\psi(\bar{y}; \bar{y_1}, \ldots, \bar{y_k})$ such that for any $\pi_1, \ldots, \pi_k \in M^{[\bar{y}]}$ if $\Delta = \psi(M; \pi_1, \ldots, \pi_k)$ then $\mathcal{I}(\Delta) = \theta_\psi(M; \pi_1, \ldots, \pi_k)$.

Observe, that every abstract cell decomposition $\mathcal{T}$ weakly definable with $\Psi(\bar{x}; \bar{y_1}, \ldots, \bar{y_k})$ induces the definable abstract cell decomposition $\mathcal{T}^\text{max}$ with

$$\mathcal{I}(\Delta) = \{ \pi \in M^{[\bar{x}]} : \varphi(\bar{x}; \pi) \text{ crosses } \Delta \}. $$

For any finite $S \in M^{[\bar{x}]}$ it holds, that $|\mathcal{T}^\text{max}(S)| = O(|S|^k)$.

5.5 Infinite model of all graphs of twin-width bounded by a fixed integer

For the rest of this section fix $\Sigma$ and $\Pi$ to be binary relational signatures $\{E, \preceq\}$ and $\{E, \preceq, \succeq\}$ respectively. We already observed in Section 5.2 that for any graph $G$ of twin-width at most $t$ we can find a $t$-proper total order on its vertices $\preceq$ (i.e. $M_\preceq(G)$ is $2t+6$-mixed free and $(G, \preceq, \succeq)$ seen as a binary $\Sigma$-structure has twin-width bounded by $t + 2$). Sometimes, by abuse of notation, we will denote $\preceq$ for a binary structure $(G, \preceq)$ by $\preceq_G$ to avoid confusion with $\preceq$ on other structures. We will do the same for the edge relation $E_G$.

For a fixed integer $t$, let $\mathcal{C}_t$ be the class of all (up to isomorphism) finite $\Sigma$-structures $(G, \preceq_G)$ such that $\preceq$ is a $t$-proper order for $G$. We may assume that for any $(G, \preceq_G) \in \mathcal{C}_t$ we have $V(G) \subseteq \mathbb{N}$ and for any distinct $(G, \preceq_G), (G', \preceq_{G'}) \in \mathcal{C}_t$ their vertex sets are disjoint. We also assume that $\mathcal{C}_t = \{(G_1, \preceq_{G_1}), (G_2, \preceq_{G_2}), \ldots\}$, i.e. graphs in $\mathcal{C}_t$ are ordered by natural numbers.
VC-density and abstract cell decomposition for edge relation in graphs of bounded twin-width

Let \( \mathbb{M}_t \) be the infinite \( \Pi \)-structure that is the disjoint sum of all structures in \( \mathcal{C}_t \) together with the total order \( \preceq \) on its universe, which is consistent with the order of graphs in \( \mathcal{C}_t \). Formally, the domain of \( \mathbb{M}_t \) is the sum of all vertex sets of structures in \( \mathcal{C}_t \) (i.e. \( \mathbb{M}_t = \bigcup_{(G, \preceq_G) \in \mathcal{C}_t} V(G) \)), and:

- for any \( x, y \in M_t \) we have \( E(x, y) \) if and only if there is a \((G, \preceq_G) \in \mathcal{C}_t \) such that \( x, y \in V(G) \) and \( E_G(x, y) \);
- for any \( x, y \in M_t \) we have \( x \preceq y \) if and only if there is a \((G, \preceq_G) \in \mathcal{C}_t \) such that \( x, y \in V(G) \) and \( x \preceq_G y \);
- for any \( x, y \in M_t \) we have \( x \succeq y \) if and only if there are \((G_i, \preceq_{G_i}) \in \mathcal{C}_t \) such that \( x \in V(G_i) \), \( y \in V(G_j) \) and either \( i < j \) or it holds that \( i = j \) and \( x \preceq_{G_i} y \).

By Theorem 3.1 for every integer \( t \) there is some constant \( k \) such that for every graph \( G \) of twin-width at most \( t \) and for all vertices \( a_i \in V(G)(i \in [l]) \) and \( b_i \in V(G)(i \in [l]) \) if

\[ E(a_i, b_j) \iff i \in I, \]

then \( l \leq k \). Indeed, the negation of this statement would contradict the fact, that classes of graphs of bounded twin-width admit linear neighborhood complexity.

Using this fact and [6, Theorem 39] we can prove the following lemma.

**Lemma 5.3.** For every integer \( t \) the structure \( \mathbb{M}_t \) is monadically NIP.

The proof of Lemma 5.3 is rather simple but technical, so we defer it to Appendix A.

Now we show another property of \( \mathbb{M}_t \) – the edge relation in this structure admits a definable abstract cell decomposition.

**Theorem 5.4.** For any integer \( t \) formula \( \varphi(x, y) \equiv E(x, y) \) admits a definable abstract cell decomposition \( T \) in the structure \( \mathbb{M}_t \), such that for any finite \( S \subset M_t \) we have \( |T(S)| = |S|^{O(t)} \).

**Proof.** We know that the formula \( \psi_t \) from Proposition 4.5 gives a weakly definable abstract cell decomposition \( T' \) of \( \varphi \) for any ordered graph \((G, \preceq) \in \mathcal{C}_t \). Therefore, for any \( G_i \) such that \( A = S \cap V(G_i) \neq \emptyset \) we can directly use \( \psi_t \) to partition \( V(G_i) \) into subsets with the same neighborhood in \( S \) using parameters from \( A \). It remains to split into definable sets the vertices of \( G_j \)’s such that \( S \cap V(G_j) = \emptyset \). This can be easily done by using pairs of elements \( s_k, s_l \in S \) such that \( s_k \succeq s_l \) and no other element from \( S \) is between them. Indeed, we just define sets of vertices that are between such pairs \( s_k, s_l \) in \( \succeq \) and are not in \( \preceq \) relation neither with \( s_k \) nor with \( s_l \).

By remark in Section 5.4, there is also a definable abstract cell decomposition \( T = T^{\max} \) weakly definable by the formulas, we described. As they have \( O(t) \) parameters, we have that for any finite \( S \subset M_t \) the condition \( |T(S)| = |S|^{O(t)} \) holds.

Lemmas 5.3 and 5.4 allow us to derive a number of combinatorial tools known for distal structures for graphs of bounded twin-width. A structure is distal if every formula \( \varphi(\vec{X}, \vec{Y}) \) admits a definable abstract cell decomposition. Although we did not show that \( \mathbb{M}_t \) is distal, we know that edge relation in \( \mathbb{M}_t \) admits definable abstract cell decomposition. That is enough to formulate these tools.

### 5.6 Cutting lemma for graphs of bounded twin-width

By simple application of [9, Theorem 3.2] to \( \mathbb{M}_t \) we obtain the following theorem.
Theorem 5.5. For any fixed integer $t$ there exists a constant $C = C(t)$, such that for any $(G, \leq) \in \mathcal{C}_t$, any $A \subseteq V(G)$ of size $n$ and any real $r$ satisfying $1 \leq r \leq n$ we can partition $V(G)$ into sets $X_1, \ldots, X_l$ with

$$l \leq C r^{O(t)},$$

such that for every $X_i$ there are at most $\frac{n}{2}$ vertices in $a \in A$ for which $N_{X_i}(a)$ crosses $X_i$ (i.e. there are vertices $x, x' \in X_i$ satisfying $E(x, a)$ and $\neg E(x', a)$). Moreover, each of the $X_i$‘s is an intersection of at most two $\psi_t(x; A)$ definable sets.

5.7 Regularity lemma for graphs of bounded twin-width

By repeating the proof of [10, Theorem 5.8] and applying it to $M_t$ we obtain a strong regularity lemma for graphs of bounded twin-width.

Theorem 5.6. For any fixed integer $t$ and any $\varepsilon > 0$ there are a constant $c = c(t)$ and $K = O(t((1/\varepsilon)^c))$ such that for any graph $G$ of twin-width at most $t$ there is a partition $P_1, \ldots, P_K$ of vertices of $G$, such that

$$\sum |P_i||P_j| \leq \varepsilon |V(G)|^2$$

where the sum is over all $i, j \in |V(G)|$ such that $P_i$ and $P_j$ are not homogeneous.

References

1. Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko. Vapnik-Chervonenkis density in some theories without the independence property, I. Transactions of the American Mathematical Society, 368:5889–5949, 2016. doi:https://doi.org/10.1090/tran/6659.
2. Patrick Assouad. Densité et dimension. Annales de l’Institut Fourier, 33:233–282, 1983.
3. J. T. Baldwin and S. Shelah. Second-order quantifiers and the complexity of theories. Notre Dame Journal of Formal Logic, 26(3):229 – 293, 1985. doi:10.1305/ndjfl/1093870870.
4. Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width III: Max Independent Set, Min Dominating Set, and Coloring. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, 48th International Colloquium on Automata, Languages, and Programming (ICALP 2021), volume 198 of Leibniz International Proceedings in Informatics (LIPIcs), pages 35:1–35:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. URL: https://drops.dagstuhl.de/opus/volltexte/2021/14104, doi:10.4230/LIPIcs.ICALP.2021.35.
5. Édouard Bonnet, Eun Jung Kim, Amadeus Reinald, Stéphan Thomassé, and Rémi Watrigant. Twin-Width and Polynomial Kernels. In Petr A. Golovach and Meirav Zehavi, editors, 16th International Symposium on Parameterized and Exact Computation (IPEC 2021), volume 214 of Leibniz International Proceedings in Informatics (LIPIcs), pages 10:1–10:16, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. URL: https://drops.dagstuhl.de/opus/volltexte/2021/15393, doi:10.4230/LIPIcs.IPEC.2021.10.
6. Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width I: tractable FO model checking. In Sandy Irani, editor, 61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020, pages 601–612. IEEE, 2020. doi:10.1109/FOCS46700.2020.00062.
7. Bernard Chazelle. Cuttings, chapter chapter26. CRC Press, 2018. URL: https://www.routledgehandbooks.com/doi/10.1201/9781315119335-26, doi:10.1201/9781315119335-26.
VC-density and abstract cell decomposition for edge relation in graphs of bounded twin-width

8 Bernard Chazelle, Herbert Edelsbrunner, Leonidas J. Guibas, and Micha Sharir. A singly exponential stratification scheme for real semi-algebraic varieties and its applications. *Theor. Comput. Sci.*, 84(1):77–105, Jul 1991. doi:10.1016/0304-3975(91)90261-Y.

9 Artem Chernikov, David Galvin, and Sergei Starchenko. Cutting lemma and zarankiewicz’s problem in distal structures, 2020. arXiv:1612.00908.

10 Artem Chernikov and Sergei Starchenko. Regularity lemma for distal structures, 2016. arXiv:1507.01482.

11 Josef Cibulka and Jan Kynčl. Füredi-Hajnal limits are typically subexponential, 2016. arXiv:1607.07491.

12 David Haussler. Sphere packing numbers for subsets of the boolean n-cube with bounded Vapnik-Chervonenkis dimension. *Journal of Combinatorial Theory, Series A*, 69(2):217–232, 1995. URL: https://www.sciencedirect.com/science/article/pii/0097316595900527, doi:https://doi.org/10.1016/0097-3165(95)90052-7.

13 Janos Komlos and Miklós Simonovits. Szemerédi’s regularity lemma and its applications in graph theory. *Combinatorica*, 2, 01 1996.

14 Michael C. Laskowski. Vapnik-Chervonenkis Classes of Definable Sets. *Journal of the London Mathematical Society*, s2-45(2):377–384, 04 1992. doi:10.1112/jlms/s2-45.2.377.

15 László Lovász and Balázs Szegedy. Regularity Partitions and The Topology of Graphons, pages 415–446. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. doi:10.1007/978-3-642-14444-8_12.

16 Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory, Series A*, 107(1):153–160, 2004. URL: https://www.sciencedirect.com/science/article/pii/S0097316504000512, doi:https://doi.org/10.1016/j.jcta.2004.04.002.

17 Jiri Matousek. *Lectures on Discrete Geometry*. Springer-Verlag, Berlin, Heidelberg, 2002.

18 Jiří Matoušek, Emo Welzl, and Lorenz Wernisch. Discrepancy and approximations for bounded VC-dimension. *Combinatorica*, 13(4):455–466, 1993. doi:https://doi.org/10.1007/BF01303517.

19 Jaroslav Nešetřil, Patrice Ossona de Mendez, Roman Rabinovich, and Sebastian Siebertz. Classes of graphs with low complexity: The case of classes with bounded linear rankwidth. *European Journal of Combinatorics*, 91:103223, 2021. Colorings and structural graph theory in context (a tribute to Xuding Zhu). URL: http://www.sciencedirect.com/science/article/pii/S019566982030144X, doi:https://doi.org/10.1016/j.ejc.2020.103223.

20 Adam Paszke and Michał Pilipczuk. VC Density of Set Systems Definable in Tree-Like Graphs. In Javier Ejarza and Daniel Král, editors, *45th International Symposium on Mathematical Foundations of Computer Science (MFCS 2020)*, volume 170 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 78:1–78:13, Dagstuhl, Germany, 2020. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. URL: https://drops.dagstuhl.de/opus/volltexte/2020/12747, doi:10.4230/LIPIcs.MFCS.2020.78.

21 Felix Reidl, Fernando Sánchez Villaamil, and Konstantinos Stavropoulos. Characterising bounded expansion by neighbourhood complexity. *European Journal of Combinatorics*, 75:152–168, 2019. URL: https://www.sciencedirect.com/science/article/pii/S0195669818301355, doi:https://doi.org/10.1016/j.ejc.2018.08.001.

22 Norbert Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972. URL: https://www.sciencedirect.com/science/article/pii/0097316572900192, doi:https://doi.org/10.1016/0097-3165(72)90019-2.
A Structure $\mathcal{M}_t$ is monadically NIP

In this appendix we prove that the binary structure $\mathcal{M}_t$, which is the ordered sum of all the ordered graphs of twin-width at most $t$, is monadically NIP and hence NIP. For the proof we state basic facts about first-order types and Ehrenfeucht-Fraïssé games. We do it based on [26].

A.1 Types and Ehrenfeucht-Fraïssé games

For a first-order formula $\varphi$ we define its quantifier rank as the depth of nesting of its quantifiers.

Let $\pi$ be a tuple of elements of a structure $\mathcal{A}$. We define its type of quantifier rank $q$ as the set of all formulas $\varphi(x)$ of quantifier rank $q$ with $|a| = |x|$ such that $\mathcal{A} \models \varphi(a)$. We denote the type of quantifier rank $q$ of $a$ by $\text{tp}^{q}(a)$. We can check that $a$ has a specific type of rank $q$ by an FO formula, which is formalised by the following lemma.

Lemma A.1. For any integer $l$ and any type $\tau$ of quantifier rank $q$ of tuples of length $l$ there is an FO formula $\theta_{q}(\pi)$ such that for any structure $\mathcal{A}$ and any tuple $\pi$ of its elements with $|\pi| = l$ we have

$$\mathcal{A} \models \theta_{q}(\pi) \iff \text{tp}^{q}(\pi) = \tau.$$  

For any integer $r$ and two tuples $\pi, \bar{b} \in A^l$ we consider $r$-round Ehrenfeucht-Fraïssé game played between two players, Spoiler and Duplicator. Each of the players has a set of $l + r$ pebbles, which will be placed on the elements of $A$. Initially, the first $l$ pebbles of Spoiler are placed on the elements $a_1, \ldots, a_l$, and the first $l$ pebbles of Duplicator are placed on the elements $b_1, \ldots, b_l$. The game then proceeds in $r$ rounds, where in each round, first Spoiler places a pebble on some element in $A$, and then Duplicator responds by doing the same thing. After $r$ rounds, if the substructures of $A$ induced by the $l + r$ pebbles of Spoiler and Duplicator are identical, then Duplicator wins, and loses otherwise.

The following theorem describes a connection between types and Ehrenfeucht-Fraïssé games.

Theorem A.2 (Ehrenfeucht-Fraïssé Theorem). Let $\pi = (a_1, \ldots, a_l)$ and $\bar{b} = (b_1, \ldots, b_l)$ be two tuples of elements of a structure $\mathcal{A}$. Duplicator has a winning strategy in $r$-round Ehrenfeucht-Fraïssé game between $\pi$ and $\bar{b}$ if and only if $\text{tp}^{r}(\pi) = \text{tp}^{r}(\bar{b})$.  


A.2 Proof of Lemma 5.3

Lemma 5.3. For every integer $t$ the structure $M_t$ is monadically NIP.

Proof. Take any expansion $\Pi^+$ of $\Pi$ by unary predicate symbols and any FO $\Pi^+$-formula $\varphi(x, y)$ of quantifier rank $q$. Let $M^+_t$ be a monadic lift of $M_t$. Assume that there are some elements $a_i (i \in \{1, \ldots, l\})$ and $b_I (I \subseteq \{1, \ldots, l\})$ of $M_t$ such that

$$M^+_t \models \varphi(a_i, b_I) \iff i \in I.$$  

These elements appear only in finite number of structures $(G, \preceq_G)$, in particular we can assume that they all appear in first $n$ graphs of $\mathcal{C}_t$. Let us denote $V_n = V(G_1) \cup \ldots \cup V(G_n)$ and $M^+_{t,n}$ be the substructure of $M^+_t$ induced by $V_n$.

Let us take any $x, y, x', y' \in V_n$ such that $tp^q(x, y) = tp^q(x', y')$ in $M^+_t$. We will show, using Ehrenfeucht-Fraïssé games, that $tp^q(x, y) = tp^q(x', y')$ also in whole $M^+_t$. Consider the following strategy for Duplicator in the $q$-round game. Whenever Spoiler places a pebble on an element in $V_n$, then Duplicator answers according to her winning strategy for the structure $M^+_{t,n}$ (she has one because of our assumption). On the other hand, if Spoiler places a pebble on an element of $M_t$ that is not in $V_n$, then Duplicator answers with the same element. After $q$ rounds elements of $V_n$ with pebbles of Spoiler and Duplicator induce the same substructure of $M^+_t$ (because of our assumption about winning strategy for Duplicator on $M^+_{t,n}$) and the same is true for the elements in $M_t \setminus V_n$ (Duplicator and Spoiler placed pebbles on the same elements). Moreover, all the elements with pebbles in $V_n$ and in $M_t \setminus V_n$ are related to each other only with $\preceq$ (and in the same way). That proves winning strategy for Duplicator. Similarly, we can prove that if $tp^q(x, y) \neq tp^q(x', y')$ in $M^+_{t,n}$, then we also have $tp^q(x, y) \neq tp^q(x', y')$ in $M^+_t$.

The argument above proves that equalities and inequalities of types of pairs of elements are preserved if we go from $M^+_t$ to its induced substructure $M^+_{t,n}$. As we can specify types of rank $q$ with formulas, there is some $\Pi^+$-formula $\psi(x, y)$ such that

$$M^+_t \models \varphi(a_i, b_I) \iff M^+_{t,n} \models \psi(a_i, b_I)$$

for all $i \in [l]$ and $I \subseteq [l]$. Finally, it is straightforward to see, that the substructure of $M_t$ induced by $V_n$ has twin-width at most $t + 2$, so by applying [6, Theorem 39] we get that $l$ is bounded by some constant depending only on $t$ and cannot be arbitrarily large. ◀