Neighbourhood complexity of graphs of bounded twin-width*

Édouard Bonnet† Florent Foucaud‡ § Tuomo Lehtilä¶ ‖ Aline Parreau**

January 12, 2023

Abstract

We give essentially tight bounds for, $\nu(d,k)$, the maximum number of distinct neighbourhoods on a set X of k vertices in a graph with twin-width at most d. Using the celebrated Marcus-Tardos theorem, two independent works [Bonnet et al., Algorithmica '22; Przybyszewski '22] have shown the upper bound $\nu(d,k) \leqslant \exp(\exp(O(d)))k$, with a double-exponential dependence in the twin-width.

We give a short self-contained proof that for every d and k,

$$\nu(d,k) \leqslant (d+2)2^{d+1}k = 2^{d+O(\log d)}k,$$

and build a bipartite graph implying $\nu(d,k) \geqslant 2^{d+\log d+O(1)}k$, in the regime when k is large enough compared to d.

1 Introduction

The aim of this paper is to refine our understanding of how complex the neighbourhoods of graphs of bounded twin-width can be. We provide an improved bound on the neighbourhood complexity of such graphs, complemented by a construction showing that our bound is essentially tight. The improvements in the bounds for neighbourhood complexities translate directly to better structural bounds and algorithms, in some contexts which are explained below.

Twin-width. Twin-width is a recently introduced graph invariant [10]; see Section 2 for a definition. It can be naturally extended to matrices over finite alphabets and binary structures [10, 7, 12]. Although classes of bounded twin-width are broad and diverse, they allow (most of the time, provided a witness is given as an input) improved algorithms, compared to what is possible on general graphs or binary structures.

Most prominently, it was shown [10] that, on n-vertex graphs given with a d-sequence (a witness that their twin-width is at most d), deciding if a first-order sentence φ holds can be solved in time $f(d,\varphi)n$, for some computable function f. In some special cases, such as for k-Independent Set or k-Dominating Set[1], single-exponential parameterised algorithms running in time $2^{O_d(k)}n$ are possible [5]. In the same setting, the triangles of an n-vertex m-edge graph can be counted in time $O(d^2n+m)$ [19]. See [8, 18, 25] for more applications of twin-width with an algorithmic flavour.

Classes of binary structures with bounded twin-width include bounded treewidth, and more generally, bounded clique-width classes, proper minor-closed classes, posets of bounded width (that is, whose antichains are of bounded size), hereditary subclasses of permutations, as well as $\Omega(\log n)$-subdivisions of

*Florent Foucaud was financed by the French government IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25) and by the ANR project GRALMECO (ANR-21-CE48-0004). Tuomo Lehtilä's research was supported by the Finnish Cultural Foundation and by the Academy of Finland grant 338797.

†Univ Lyon, CNRS, ENS de Lyon, Université Claude Bernard Lyon 1, LIP UMR5668, France.

‡Université Clermont-Auvergne, CNRS, Mines de Saint-Étienne, Clermont-Auvergne-INP, LIMOS, 63000 Clermont-Ferrand, France.

§Univ. Orléans, INSA Centre Val de Loire, LIFO EA 4022, F-45067 Orléans Cedex 2, France.

¶Univ Lyon, UCBL, CNRS, LIRIS - UMR 5205, F69622, France.

‖University of Turku, Department of Mathematics and Statistics, Turku, Finland.

**Univ Lyon, CNRS, INSA Lyon, UCBL, Centrale Lyon, Univ Lyon 2, LIRIS, UMR5205, F-69622 Villeurbanne, France.

[1]That is, the problems of deciding whether in an input graph, there are k vertices that are pairwise non-adjacent or whose closed neighbourhood is the entire vertex set, respectively.

n-vertex graphs [10], and particular classes of (bounded-degree) expanders [6]. A rich range of geometric graph classes have bounded twin-width such as map graphs, bounded-degree string graphs [10], classes with bounded queue number or bounded stack number [6], segment graphs with no $K_{t,t}$ subgraph, and visibility graphs of simple polygons without large independent sets [4], to give a few examples.

If efficiently approximating the twin-width is a challenging open question in general, this is known to be possible for the above-mentioned classes (albeit a representation may be needed for the geometric classes) and for ordered graphs [7]. By that, we mean that there are two computable functions f, g and an algorithm that, for an input n-vertex graph G from the class and an integer k, and in time $g(k)n^{O(1)}$, either outputs an $f(k)$-sequence (again, witnessing that the twin-width is at most $f(k)$) or correctly reports that the twin-width of G is larger than k.

Structural properties of graph classes of bounded twin-width include χ-boundedness [5], even with a quasipolynomial binding function [24], smallness (i.e., containing up to isomorphism $2^{O(n)}$ n-vertex graphs) [6, 12], and Vapnik-Chervonenkis (VC) density at most 1 [9, 26]. The latter property is the topic of the current article.

VC density and neighbourhood complexity. VC density is related to the celebrated VC dimension [29]. Given a set-system (or hypergraph) $\mathcal{S}$ on a domain X, the *shatter function* $\pi_{\mathcal{S}} : \mathbb{N} \to \mathbb{N}$ is defined as

$$\pi_{\mathcal{S}}(n) = \max_{A \in \binom{X}{n}} |\{Y \subseteq A \mid \exists S \in \mathcal{S}, Y = A \cap S\}|.$$

The Perles-Sauer-Shelah lemma states that $\pi_{\mathcal{S}}(n) = O(n^d)$ if the VC dimension of $\mathcal{S}$ (i.e., the supremum of $\{n \mid \pi_{\mathcal{S}}(n) = 2^n\}$) is a finite integer d. Then the VC density of $\mathcal{S}$ is defined as $\inf\{c \in \mathbb{R} \mid \pi_{\mathcal{S}}(n) = O(n^c)\}$, and as $+\infty$ if the VC dimension is unbounded.

We define the *VC density* of an infinite class $\mathcal{C}$ of finite graphs as the VC density of the infinite set-system formed by the neighbourhood hypergraph of the disjoint union of the graphs of $\mathcal{C}$, that is, $\{N_G(v) \mid v \in V(\uplus_{G \in \mathcal{C}} G)\}$, where $N_G(v)$ denotes the set of neighbours of v in G. The VC density is an important measure in finite model theory, often more tractable than the VC dimension (see for instance [1, 2]). Tight bounds have been obtained for the VC density of (logically) definable hypergraphs from graph classes of bounded clique-width [23] (with monadic second-order logic), and more recently, of bounded twin-width [18] (with first-order logic).

In structural graph theory and kernelisation [16] (a subarea of parameterised complexity [14]) the function $\pi_{\mathcal{N}(G)}$, where $\mathcal{N}(G)$ is the neighbourhood hypergraph of G, is often[1] called *neighbourhood complexity*. (See [3] for an algorithmic study of the computation of this notion.) In these contexts, obtaining the best possible upper bound for $\pi_{\mathcal{N}(G)}$ (and not just the exponent matching the VC density) translates to qualitatively better structural bounds and algorithms; see for instance [9, 11, 15, 28].

The *r-neighbourhood complexity* of G is the neighbourhood complexity of G^r, with same vertex set as G, and an edge between two vertices at distance at most r in G. Reidl et al. [28] showed that among subgraph-closed classes, bounded expansion[2] is equivalent to linear r-neighbourhood complexity. Indeed, the more general nowhere dense classes [21] (another invention of the Sparsity program [22]) have almost linear r-neighbourhood complexity [15]: there is a function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that for every $\varepsilon > 0$, $\pi_{\mathcal{N}(G^r)}(n) \leqslant f(r, \varepsilon)n^{1+\varepsilon}$ for all n. On hereditary classes, i.e., closed under taking *induced* subgraphs, there is no known characterisation of linear neighbourhood complexity.

As we already mentioned in a different language, bounded twin-width classes have been proven to have linear neighbourhood complexity. See [9, Lemma 3] or [26, Section 3] for two independent proofs, both using the Marcus-Tardos theorem [20]. However, the dependence in the twin-width is doubly exponential in both papers. Setting $\nu(d, k)$ as the maximum number of distinct neighbourhoods on a set of size k within a graph of twin-width at most d, i.e., $\max\{\pi_{\mathcal{N}(G)}(k) \mid G \text{ has twin-width at most } d\}$, they show that $\nu(d, k) \leqslant \exp(\exp(O(d)))k$.

Our results. In this note, we give in Section 3 a self-contained proof (not using the Marcus-Tardos theorem) that $\nu(d, k) \leqslant 2^{d+O(\log d)}k$. In Section 4, we complement that proof with a construction of a

[1] Some authors define the neighbourhood complexity as $n \mapsto \frac{\pi_{\mathcal{N}(G)}(n)}{n}$.

[2] A notion from the Sparsity theory of Nešetřil and Ossona de Mendez [22] extending bounded degree and proper minor-free classes.

bipartite graph witnessing that $\nu(d,k) \geqslant 2^{d+\log d+O(1)}k$, which makes our single-exponential upper bound in twin-width essentially tight.

2 Preliminaries

We use the standard graph-theoretic notations: $V(G)$, $E(G)$, $G[S]$, $G-S$ respectively denote the vertex set, edge set, subgraph of G induced by S, and subgraph of G induced by $V(G) \setminus S$. If $v \in V(G)$, then $N_G(v)$ (or $N(v)$ if G is clear from the context) denotes the set of neighbours of v in G. If $X \subseteq V(G)$, then an *X-neighbourhood* is a set $N(v) \cap X$ for some $v \in V(G)$.

We now define the twin-width of a graph, following the definition of [10].

A *trigraph* is a triple $G = (V(G), E(G), R(G))$ where $E(G)$ and $R(G)$ are two disjoint sets of edges on $V(G)$: the usual edges (also called *black edges*) and the *red edges*. Informally, a red edge between two vertices u and v means that some errors have been made between u and v. The *red degree* of a trigraph is the maximum degree of the graph $(V(G), R(G))$. Any graph G can be interepreted as a trigraph $G = (V(G), E(G), \emptyset)$. Given a trigraph and two vertices $u, v \in V(G)$ (not necessarily adjacent), the trigraph $G/u,v = G'$ is obtained by *contracting* u and v in a new vertex w such that:

- $V(G') = \{w\} \cup V(G) \setminus \{u,v\}$;
- the edges between vertices of $V(G) \setminus \{u,v\}$ are the same in G';
- the following edges are incident to w:
  - $wx \in E(G')$ if $xu \in E(G)$ and $xv \in E(G)$;
  - $wx \notin E(G') \cup R(G')$ if $xu \notin E(G) \cup R(G)$ and $xv \notin E(G) \cup R(G)$;
  - $wx \in R(G')$ otherwise.

In other words, the common black neighbours of u and v are black neighbours of w. All the other neighbours of u or v are red neighbours of w. Red edges stay red, black edges stay black, red and black edges become red. We say that $G/u,v$ is a *contraction* of G. A *d-sequence* of an n-vertex graph G is a sequence of n trigraphs $G_n = G, G_{n-1}, ..., G_1$ such that each trigraph G_i is obtained from G_{i+1} by a contraction and has red degree at most d. The *twin-width* of G, denoted by $\mathrm{tww}(G)$, is the minimum integer d such that G admits a d-sequence. Note that an induced subgraph of G has a twin-width smaller or equal to the twin-width of G [10].

If $u \in G_i$, then $u(G)$ denotes the set of vertices of G eventually contracted to u in G_i. Instead of considering the trigraphs G_i, we might prefer to deal with the partitions induced by the sets $u(G)$ for u in G_i: $\mathcal{P}_i = \{u(G) \mid u \in V(G_i)\}$. We say that there is a red edge between two parts $u(G)$ and $v(G)$ of $\mathcal{P}_i$ if uv is red in G_i.

3 Upper bound on the number of distinct neighbourhoods

We state and prove our upper bound on the maximum number of distinct X-neighbourhoods in bounded twin-width graphs.

Theorem 1. *Let G be an n-vertex graph of twin-width d, and $X \subseteq V(G)$. Then the number of distinct X-neighbourhoods in G is at most $(d+2)2^{d+1}|X| = 2^{d+O(\log d)}|X|$.*

Proof. Fix $X \subseteq V(G)$. First of all, for all vertices of $V(G) \setminus X$ with the same X-neighbourhood, we keep only one representative. Note that the new graph G'' is an induced subgraph of G, thus its twin-width is at most d. We further modify graph G'' by adding for each $v \in X$ a new vertex u to G'' so that $N(u) = N(v)$ if such vertex does not exist in $V(G'') \setminus X$. The new graph is called G' and it has the same twin-width as G''.

Let $M = (d+2)2^{d+1} + 1$. We prove by induction on n that an n-vertex graph of twin-width at most d with a set X of k vertices, where all vertices outside X have a distinct X-neighbourhood, satisfies $n \leqslant kM$. This will prove that G' has at most kM vertices, and thus that in G, there are at most $(M-1)k$ distinct X-neighbourhoods.

The statement is trivially true for $n \leqslant 5$ since $M \geqslant 5$, for all $d \geqslant 0$.

Thus, assume $n \geqslant 6$. In particular, we have $k > 1$. Let $x \in X$. Let $X' = X \setminus \{x\}$ and let T_x be the set of pairs of vertices outside X that are twins with respect to X', i.e.

$$T_x = \left\{ \{u, v\} \in \binom{V(G') \setminus X}{2} \mid N(u) \cap X' = N(v) \cap X' \right\}.$$

Since every vertex of $V(G') \setminus X$ has a distinct neighbourhood in X, there are at most two vertices of $V(G') \setminus X$ with the same (possibly empty) neighbourhood N in X'; namely the vertices $u, v \in V(G') \setminus X$ with $N(u) \cap X = N$ and $N(v) \cap X = N \cup \{x\}$ (if they exist). Hence, T_x consists of pairwise-disjoint pairs of vertices.

We prove the following claim.

Claim A. *There exists a vertex x of X such that T_x comprises at most $M - 1$ pairs, in G'.*

Proof of claim. By contradiction, assume this is not the case: for every x in X, T_x has size at least M. Consider a d-sequence of contractions $G'_n, \ldots, G'_1$ of G'. Consider the last step G'_i of the sequence where all the parts of $\mathcal{P}_i$ contain at most one vertex of X (that is, contrary to $\mathcal{P}_i$, some part of $\mathcal{P}_{i-1}$ contains two vertices of X).

Let P be a part of $\mathcal{P}_i$. Let x be the unique (if there exists one) element of $P \cap X$. Then we claim that $|P \setminus X| \leqslant 2^{d+1}$. Indeed, any two vertices of $P \setminus X$ have some vertex in the symmetric difference of their X-neighbourhoods, either it is x, or some vertex x' of X outside P. If that distinguishing vertex is some x' that is not in P, then there has to be a red edge between P and the part that contains x'. There are at most d red edges with P as an extremity. Since all the elements of X are in distinct parts in G'_i, it means that $d + 1$ vertices of X are enough to distinguish all the X-neighbourhoods of vertices of $P \setminus X$, and thus $|P \setminus X| \leqslant 2^{d+1}$.

We now consider the next contraction in the sequence, which leads to G'_{i-1}. By definition of G'_i, it must contract two vertices corresponding to two parts of $\mathcal{P}_i$ that both contain an element of X. Let x_1 and x_2 be these two elements of X. Let Q be the part of $\mathcal{P}_{i-1}$ that contains both x_1 and x_2. By our assumption, T_{x_1} has size at least M. Let $\{u, v\}$ be a pair of T_{x_1}. Since u and v have the same neighbourhood in $X \setminus \{x_1\}$, it means that they are either both adjacent or both non-adjacent to x_2, and exactly one of them is adjacent to x_1. Thus, necessarily, one vertex among the pair $\{u, v\}$ is adjacent to exactly one vertex among $\{x_1, x_2\}$. In particular, if this vertex is not in Q, then there has to be a red edge between the part containing this vertex and the part Q in G'_{i-1}. Since T_{x_1} contains at least M pairs (which are disjoint) and Q has at most 2^{d+2} vertices not in X, there are at least $M - 2^{d+2}$ vertices not in X whose part in G'_{i-1} has a red edge to Q. Since each other part has at most 2^{d+1} vertices not in X, it makes at least $\frac{M-2^{d+2}}{2^{d+1}}$ red edges incident to Q. Thus, we must have $\frac{M-2^{d+2}}{2^{d+1}} \leqslant d$, leading to $M \leqslant 2^{d+1}(d + 2)$, a contradiction that proves the claim. (□)

By Claim A, there exists a vertex $x \in X$ such that $|T_x| \leqslant M - 1$. Let Y be a set of $|T_x|$ vertices that intersects each pair of T_x exactly once. Let $G_Y = G' - (Y \cup \{x\})$. Then, $X' = X \setminus \{x\}$ is a vertex set of size $k - 1$ such that all X'-neighbourhoods of vertices outside X' are distinct. The graph G_Y has at least $n - M$ vertices, and twin-width at most d. By induction, we have $n - M \leqslant |V(G_Y)| \leqslant (k - 1)M$ and thus, $n \leqslant kM$. Hence, once we recall that no vertex in X has unique X-neighbourhood, there are at most $(M - 1)k$ distinct X-neighbourhoods, which completes the proof. □

4 Lower bound on the number of distinct neighbourhoods

Notice that when $|X|$ and $\mathrm{tww}(G)$ are roughly the same, the bound from Theorem 1 cannot be sharp, since G' has at most $2^{|X|} + |X|$ vertices. However, when $|X|$ is large enough compared to $\mathrm{tww}(G)$, we next show that the bound is sharp up to a constant factor.

Proposition 2. *There is a positive constant c, such that for any integer d, there is a bipartite graph G of twin-width at most d, and a large enough set $X \subseteq V(G)$, with at least $c \cdot d2^d|X|$ distinct X-neighbourhoods in G.*

Proof. Observe that the claim is clearly true for any small d. Thus, we do not need to consider separately graphs with small twin-width upper bounded by a constant. Hence, we assume from now on that $d \geq d'$ where d' is some positive constant.

We construct the graph G as follows. Let A, B, $C \in \mathbb{Z}$ be three constants that will be given later (A and B will be roughly equal to $\sqrt{d}$ and C will be roughly equal to d). Let $X = \{x_1, ..., x_k\}$ be an independent set of k vertices. Our goal is that each vertex in $V(G) \setminus X$ has a unique X-neighbourhood. For any integers i, j, t with $1 \leqslant i \leqslant j \leqslant i + A - 1$, $j + 2 \leqslant t \leqslant j + 1 + B$ and $t \leqslant k - C$, we create a set $V_{i,j,t}$ of vertices as follows. Consider the set $X_t = \{x_{t+1}, ..., x_{t+C}\}$. For every subset Y of X_t, let $Y' = \{x_i, ..., x_j, x_t\} \cup Y$ and add a vertex $v_{Y'}$ to $V_{i,j,t}$, making it adjacent to the vertices of Y'. Each set $V_{i,j,t}$ has size 2^C and there are $\Theta(kAB)$ (for fixed A and B and growing k) such sets. Thus there are $\Theta(kAB2^C)$ vertices in the graph.

Any two vertices not in X have distinct X-neighbourhoods. Indeed, by considering the natural ordering of X induced by the indices, any vertex not in X is first adjacent to a consecutive interval of vertices from x_i to x_j, then is not adjacent to vertices from x_{j+1} to x_{t-1} (which is not empty since $t \geqslant j + 2$), and then adjacent to x_t. Thus, if two vertices have the same X-neighbourhood, they must be in the same set $V_{i,j,t}$. But then, they have a distinct neighbourhood in $\{x_{t+1}, ..., x_{t+C}\}$.

We now prove that the twin-width of G is at most $M = \max\{AB, C\} + 2$. For that, we give a sequence of contractions with red degree at most M.

The contraction sequence is split into $k - C$ steps, for each vertex of X. Let $0 \leq i \leq k - C - 1$. Step 0 corresponds to the starting point, where each vertex is alone. Let $i \geqslant 1$. After Step i, there will be the following parts in the corresponding partition (vertices not in any part have not yet been contracted):

- For each j, t such that $i \leqslant j \leqslant i + A - 1$ and $j + 2 \leqslant t \leqslant j + 1 + B$, there is a part $B_{j,t}$. The parts $B_{i,t}$ (parts with $j = i$), contain all the vertices of the sets $V_{i',j',t}$ such that $j' \leq i$. The parts $B_{j,t}$ with $j > i$ contain all the vertices of the sets $V_{i',j',t}$ such that $i' \leqslant i$ and $j' = j$. Note that there are AB non-empty $B_{j,t}$ parts in total.
- There is a part X_0 that contains vertices from x_1 to x_i of X.
- There is a part T (for "trash") that contains all the vertices of the sets $V_{i',j,t}$ with $t \leqslant i + 1$.

All the other vertices are not yet contracted. This corresponds to the vertices from x_{i+1} to x_k of X and to the vertices of the sets $V_{i',j,t}$ with $i' > i$. Indeed, if $i' \leqslant i$ and $t \leqslant i + 1$, then the vertices of $V_{i',j,t}$ are in T. If $t \geqslant i + 2$ but $j \leqslant i$, then they are in the part $B_{i,t}$. If $j > i$, then they are in the part $B_{j,t}$.

We first prove that the red degree after Step i is at most M. Then, we explain how to get from Step i to Step $i + 1$ by keeping the red degree at most M.

Consider the part $B_{j,t}$ at the end of Step i. A vertex in this part belongs to some set $V_{i',j',t}$ with $i' \leqslant i$ and $j' = j$ if $j > i$ or $j' \leqslant i$ otherwise. In particular, two vertices of $B_{j,t}$ are adjacent to all the vertices between x_{i+1} and x_j, to no vertex between x_{j+1} and x_{t-1}, to x_t, and to no vertex after x_{t+C}. Thus, there is a red edge between the parts $B_{j,t}$ and X_0, and C red edges between the part $B_{j,t}$ and the vertices $\{x_{t+1}, ..., x_{t+C}\}$. Therefore, the number of red edges incident with $B_{j,t}$ is at most $C + 1$.

Consider now the part T. Vertices in T are adjacent only to vertices of X up to x_{i+C+1}. Since vertices x_1 to x_i are all in the part X_0, the red degree of T is at most $C + 2$.

Single vertices not in X have no incident red edges: indeed, they are all in some sets $V_{i',j,t}$ for $i' > i$ and thus are not adjacent to any vertex of X_0. For the same reason, there are red edges incident to X_0 only to T and to the parts $B_{j,t}$. Hence, the red degree of X_0 is at most $AB + 1$. Similarly, the red degree of $x_{i'}$, $i' > i + 1$ is at most $AB + 1$. Moreover, the red degree of x_{i+1} is at most one. Indeed, the only red edge is between x_{i+1} and T.

Finally, the red degree after step i is at most $\max\{AB + 1, C + 2\} \leqslant M$.

Let $i \geq 0$. We now explain how we perform the contractions to go from step i to step $i + 1$.

1. (only if $i \geq 1$) For any $i + 3 \leqslant t \leqslant i + 2 + B$, merge the part $B_{i,t}$ with the part $B_{i+1,t}$. The only new red edge this merging may lead to, when $B_{i,t}$ is non-empty, is between $B_{i+1,t}$ and x_{i+1}. Thus, we add only one red edge between x_{i+1} and $B_{i+1,t}$. Thus, the red degree of $B_{i+1,t}$ is at most $C + 2$ and the red degree of x_{i+1} is at most 2.
2. Add all the vertices of $V_{i+1,j,t}$ for some j, t to the part (that might be empty at this point) $B_{j,t}$. The red degree of $B_{j,t}$ is at most $C + 2$ since we might have a red edge between $B_{j,t}$ and x_{i+1}. The number of nonempty parts $B_{j,t}$ at this point is $AB + 1$ (there is still the part $B_{i,i+2}$). Adding T, this gives $AB + 2$ red edges incident to a vertex in X (or from part X_0).

3. Add x_{i+1} to X_0. The part X_0 has red edges only to parts $B_{i+1,t}$, to $B_{i,i+2}$ and to T, but no edges to the single vertices. Thus, it has red degree at most $AB + 2$.

4. Put the part $B_{i,i+2}$ into T. This part is only adjacent to vertices up to x_{i+2+C}, and thus has $C + 2$ red edges.

Thus, at each point, the red degree is always at most $M = \max\{AB, C\} + 2$.

The process ends at step $i = k - C - 1$. Then, all the vertices not in X are in some parts, and there are at most $AB + 1$ such parts. On the other side of bipartition, we have part X_0 and $C + 1$ single vertices. Thus, the graph is bipartite with both sides of size at most M. One can contract each part independently to finish the contraction sequence.

To conclude, taking $C = d - 2$ and $A = B = \lfloor\sqrt{d-2}\rfloor$, we have $M \leqslant d$ and $kAB2^C = \Theta(kd2^d)$. Notice that we may assume that A, B and C are positive since $d \geq d'$ where d' was some well chosen positive constant. This concludes the proof. □

5 Conclusion

We have given an improved and tight upper bound for the neighbourhood complexity of graphs of bounded twin-width. Unlike the previously known (weaker) bounds, our method is simple and avoids the use of the Marcus-Tardos theorem. We hope that it can inspire future works in the area.

It is known that the twin-width of G^r can be upper-bounded by a function of the twin-width of G and r [10]. Thus, graphs of twin-width at most d have linear r-neighbourhood complexity. We leave as an interesting open problem to obtain an essentially tight twin-width dependence for the r-neighbourhood complexity.

We remark that the neighbourhood complexity is also related to *identification problems* on graphs such as *identifying codes* or *locating-dominating sets*, where one seeks a (small) set A of vertices of a graph such that all other vertices have a distinct neighbourhood in A [17]. Some works in this area about specific graph classes, are equivalent to the study of the neighbourhood complexity of these graph classes: see for example [13, 17, 27]. Moreover, we note that for graph classes with VC density 1, since any solution has linear size, the natural minimisation versions of the above identification problems have a polynomial-time constant-factor approximation algorithm (trivially select the whole vertex set), while such an algorithm is unlikely to exist in the general case [13]. Thus, our work implies a better approximation ratio for these problems, when restricted to input graph classes of bounded twin-width.

References

[1] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko. Vapnik–chervonenkis density in some theories without the independence property, II. *Notre Dame Journal of Formal Logic*, 54(3-4):311–363, 2013.

[2] 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(8):5889–5949, 2016.

[3] Cristina Bazgan, Florent Foucaud, and Florian Sikora. Parameterized and approximation complexity of Partial VC Dimension. *Theor. Comput. Sci.*, 766:1–15, 2019.

[4] Édouard Bonnet, Dibyayan Chakraborty, Eun Jung Kim, Noleen Köhler, Raul Lopes, and Stéphan Thomassé. Twin-width VIII: delineation and win-wins. In Holger Dell and Jesper Nederlof, editors, *17th International Symposium on Parameterized and Exact Computation, IPEC 2022, September 7-9, 2022, Potsdam, Germany*, volume 249 of *LIPIcs*, pages 9:1–9:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[5] 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, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPIcs*, pages 35:1–35:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[6] Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width II: small classes. *Combinatorial Theory*, 2 (2), 2022.

[7] Édouard Bonnet, Ugo Giocanti, Patrice Ossona de Mendez, Pierre Simon, Stéphan Thomassé, and Szymon Toruńczyk. Twin-width IV: ordered graphs and matrices. In Stefano Leonardi and Anupam Gupta, editors, *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 924–937. ACM, 2022.

[8] Édouard Bonnet, Eun Jung Kim, Amadeus Reinald, and Stéphan Thomassé. Twin-width VI: the lens of contraction sequences. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1036–1056. SIAM, 2022.

[9] Édouard Bonnet, Eun Jung Kim, Amadeus Reinald, Stéphan Thomassé, and Rémi Watrigant. Twin-width and polynomial kernels. *Algorithmica*, 84:3300–3337, 2022.

[10] Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width I: tractable FO model checking. *J. ACM*, 69(1):3:1–3:46, 2022.

[11] Édouard Bonnet, O-joung Kwon, and David R. Wood. Reduced bandwidth: a qualitative strengthening of twin-width in minor-closed classes (and beyond). *CoRR*, abs/2202.11858, 2022.

[12] Édouard Bonnet, Jaroslav Nešetřil, Patrice Ossona de Mendez, Sebastian Siebertz, and Stéphan Thomassé. Twin-width and permutations. *CoRR*, abs/2102.06880, 2021.

[13] Nicolas Bousquet, Aurélie Lagoutte, Zhentao Li, Aline Parreau, and Stéphan Thomassé. Identifying codes in hereditary classes of graphs and VC-dimension. *SIAM J. Discret. Math.*, 29(4):2047–2064, 2015.

[14] Marek Cygan, Fedor V Fomin, Łukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michał Pilipczuk, and Saket Saurabh. *Parameterized algorithms*, volume 4. Springer, 2015.

[15] Kord Eickmeyer, Archontia C. Giannopoulou, Stephan Kreutzer, O-joung Kwon, Michał Pilipczuk, Roman Rabinovich, and Sebastian Siebertz. Neighborhood complexity and kernelization for nowhere dense classes of graphs. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, volume 80 of *LIPIcs*, pages 63:1–63:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[16] Fedor V Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Kernelization: theory of parameterized preprocessing*. Cambridge University Press, 2019.

[17] Florent Foucaud, George B. Mertzios, Reza Naserasr, Aline Parreau, and Petru Valicov. Identification, location-domination and metric dimension on interval and permutation graphs. I. bounds. *Theor. Comput. Sci.*, 668:43–58, 2017.

[18] Jakub Gajarský, Michał Pilipczuk, Wojciech Przybyszewski, and Szymon Toruńczyk. Twin-width and types. In Mikolaj Bojanczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022, Paris, France*, volume 229 of *LIPIcs*, pages 123:1–123:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[19] Stefan Kratsch, Florian Nelles, and Alexandre Simon. On triangle counting parameterized by twin-width. *CoRR*, abs/2202.06708, 2022.

[20] Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley-Wilf conjecture. *J. Comb. Theory, Ser. A*, 107(1):153–160, 2004.

[21] Jaroslav Nešetřil and Patrice Ossona de Mendez. On nowhere dense graphs. *Eur. J. Comb.*, 32(4):600–617, 2011.

[22] Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity - Graphs, Structures, and Algorithms*, volume 28 of *Algorithms and combinatorics*. Springer, 2012.

[23] Adam Paszke and Michał Pilipczuk. VC density of set systems definable in tree-like graphs. In Javier Esparza and Daniel Král', editors, *45th International Symposium on Mathematical Foundations of Computer Science, MFCS 2020, August 24-28, 2020, Prague, Czech Republic*, volume 170 of *LIPIcs*, pages 78:1–78:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

[24] Michał Pilipczuk and Marek Sokołowski. Graphs of bounded twin-width are quasi-polynomially χ-bounded. *CoRR*, abs/2202.07608, 2022.

[25] Michał Pilipczuk, Marek Sokołowski, and Anna Zych-Pawlewicz. Compact representation for matrices of bounded twin-width. In Petra Berenbrink and Benjamin Monmege, editors, *39th International Symposium on Theoretical Aspects of Computer Science, STACS 2022, March 15-18, 2022, Marseille, France (Virtual Conference)*, volume 219 of *LIPIcs*, pages 52:1–52:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[26] Wojciech Przybyszewski. VC-density and abstract cell decomposition for edge relation in graphs of bounded twin-width. *CoRR*, abs/2202.04006, 2022.

[27] Douglas Rall and Peter J. Slater. On location-domination numbers for certain classes of graphs. *Congressus Numerantium*, 45:97–106, 1984.

[28] Felix Reidl, Fernando Sánchez Villaamil, and Konstantinos S. Stavropoulos. Characterising bounded expansion by neighbourhood complexity. *Eur. J. Comb.*, 75:152–168, 2019.

[29] Vladimir N. Vapnik and Alexey Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. In *Measures of complexity*, pages 11–30. Springer, 2015.