Twin-width and permutations

Édouard Bonnet
Univ Lyon, CNRS, ENS de Lyon, Université Claude Bernard Lyon 1, LIP UMR5668, France

Jaroslav Nešetřil
Computer Science Institute of Charles University (IUUK), Praha, Czech Republic

Patrice Ossona de Mendez
Centre d'Analyse et de Mathématique Sociales CNRS UMR 8557, France
and Computer Science Institute of Charles University (IUUK), Praha, Czech Republic

Sebastian Siebertz
University of Bremen, Bremen, Germany

Stéphan Thomassé
Univ Lyon, CNRS, ENS de Lyon, Université Claude Bernard Lyon 1, LIP UMR5668, France

Abstract

Inspired by a width invariant defined on permutations by Guillemot and Marx, the twin-width invariant has been recently introduced by Bonnet, Kim, Thomassé, and Watrigant. We prove that a class of binary relational structures (that is: edge-colored partially directed graphs) has bounded twin-width if and only if it is a first-order transduction of a proper permutation class. As a by-product, it shows that every class with bounded twin-width contains at most $2^{O(n)}$ pairwise non-isomorphic n-vertex graphs.

2012 ACM Subject Classification Theory of computation → Finite Model Theory; Mathematics of computing → Graph theory

Keywords and phrases Twin-width, first-order transductions, structural graph theory

Funding This paper is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810115 – Dynasnet).

1 Introduction

Many constructions of graphs with good structural and algorithmic properties are based on trees. In turn this often leads to important parameters and to a hierarchy of graph properties (and graph classes), which allow to treat hard problems in a parametric way. Examples include tree-depth, treewidth, cliquewidth, shrub-depth to name just a few. These notions have relevance and in many cases are a principal tool for important complexity and algorithmic results: Let us mention at least Rossman's homomorphism preservation theorem [18], Courcelle's theorem on MSO definable properties [8, 9], Robertson and Seymour's graph minor theory [17], and low tree-depth decompositions of classes with bounded expansion [14].

In this paper we consider the twin-width graph parameter, defined by Bonnet, Kim, Thomassé and Watrigant [6] as a generalization of a width invariant for classes of permutations defined by Guillemot and Marx [11]. This parameter was intensively studied recently in the context of many structural and algorithmic questions such as FPT model checking [6], graph enumeration [4], graph coloring [3], and matrices and ordered graphs [5]. We show that twin-width can be concisely expressed by special structures called here twin models. Twin models are rooted trees augmented by a set of transversal edges that satisfies two simple properties: minimality and consistency. These properties imply that every twin model admits rankings, from which we can compute a width. The twin-width of a graph then coincides with the optimal width of ranked twin-model of the graph. While this connection is technical, twin models provide a simple way to handle classes with bounded twin-width.

We show that a class $\mathscr{C}$ of binary relational structures has bounded twin-width if and only if $\mathscr{C}$ is a transduction of a proper permutation class (i.e., a class of permutations closed under sub-permutations, which avoids at least one pattern). The involved transduction being k-bounded, it is then a consequence of [13] that any class of relational structures with bounded twin-width contains at most c^n non-isomorphic structures with n vertices hence is small (i.e., contains at most $c^n\,n!$ labeled structures with n vertices). This extends the main result of [4] while not using the "versatile contraction" machinery (but only the preservation of bounded twin-width by transduction proved in [6]). This also extends a similar property for proper minor-closed classes of graphs, which can be derived from the boundedness of book thickness, as noticed by Colin McDiarmid (see the concluding remarks of [2]). This suggests that every small hereditary class of graphs contains at most c^n non-isomorphic graphs with n vertices (see Conjecture 8.1).

This paper is a combination of model-theoretic tools (relational structures, interpretations, transductions), structural graph theory and theory of permutations. The fact that any class of graphs with bounded twin-width is just a transduction of a proper permutation class is surprising at first glance and it nicely complements another model theoretic characterization of these classes: a class of graphs has bounded twin-width if and only if it is the reduct of a dependent class of ordered graphs [5]. On the other hand twin-models are interesting objects per se and in a way presents one of the weakest forms of width parameters related to trees. Note also that for other classes of sparse structures we have no such a concrete model.

We outline the proof of our main result, by relying on Figure 1. The relevant terminology will be formally introduced in the appropriate sections.

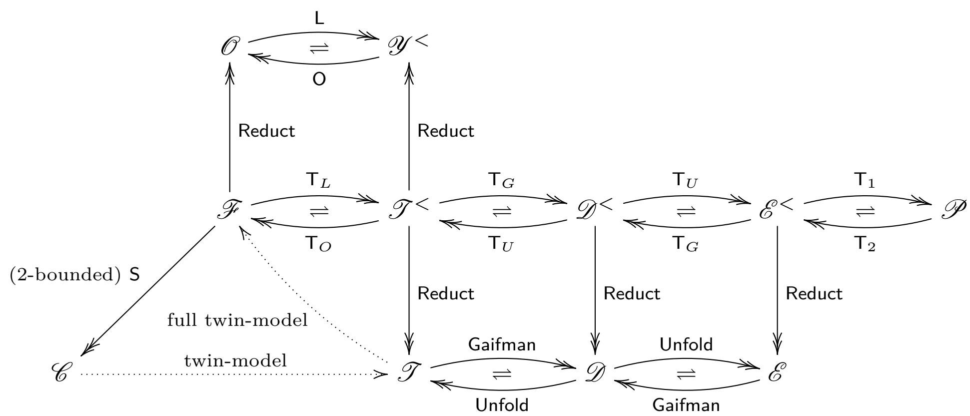

Figure 1 Relations between the classes of structures involved in the proof of the main result.

We start with a class $\mathscr{C}$ of binary relational structures with bounded twin-width. We derive a class $\mathscr{T}$ of twin-models (tree-like representations of the graphs using rooted binary trees and transversal binary relations). Replacing the rooted binary trees of the twin-models by binary tree-orders, we get a class $\mathscr{F}$ of so-called full twin-models, which we prove has bounded twin-width. This class can be used to retrieve $\mathscr{C}$ as a transduction, that is by means of a logical encoding. Using a transduction pairing (generalizing the notion of a bijective encoding) between binary tree-orders ($\mathscr{O}$) and rooted binary trees ordered by a preorder ($\mathscr{Y}^<$) we derive a transduction pairing of $\mathscr{F}$ with a class $\mathscr{T}^<$ of ordered twin-models. From the property that the class $\mathscr{D}$ of the Gaifman graphs of the twin-models in $\mathscr{T}$ is degenerate (and has bounded twin-width) we prove a transduction pairing of $\mathscr{T}$ and $\mathscr{D}$, from which we derive a transduction pairing of $\mathscr{T}^<$ and the class $\mathscr{D}^<$ of ordered Gaifman graphs of the ordered twin-models. From a transduction pairing of $\mathscr{D}$ with a class $\mathscr{E}$ of binary structures, in which each binary relation induces a pseudoforest, we deduce a transduction pairing of $\mathscr{D}^<$ with an expansion $\mathscr{E}^<$ of $\mathscr{E}$ by a linear order. We then prove a transduction pairing of this class with a class $\mathscr{P}$ of permutations. As this class has bounded twin-width (as it is a transduction of a class with bounded twin-width) we infer that $\mathscr{P}$ avoids a least one pattern. Following the backward transductions, we eventually deduce that $\mathscr{C}$ is a transduction of the hereditary closure of $\mathscr{P}$, which is a proper permutation class.

2 Preliminaries

2.1 Relational structures

We assume basic knowledge of first-order logic and refer to [12] for extensive background. A *relational signature* Σ is a finite set of relation symbols R_i with associated arity r_i. A *relational structure* $\mathbf{A}$ with signature Σ, or simply a Σ-*structure* consists of a *domain* A together with relations $R_i(\mathbf{A}) \subseteq A^{r_i}$ for each relation symbol $R_i \in \Sigma$ with arity r_i. The relation $R_i(\mathbf{A})$ is called the *interpretation* of R_i in $\mathbf{A}$. We will often speak of a relation instead of a relation symbol when there is no ambiguity. We may write $\mathbf{A}$ as $(A, R_1(\mathbf{A}), \dots, R_s(\mathbf{A}))$. In this paper we will consider relational structures with finite domain, and (mostly) with relations of arity at most 2. Without loss of generality we assume that all structures contain at least two elements. We will further assume that Σ-structures are *irreflexive*, that is, $(v, v) \notin R_i(\mathbf{A})$ for every element $v \in A$ and relation symbol $R_i \in \Sigma$. A unary relation is called a *mark*. Let R be a binary relation symbol and let $u, v \in A$. That the pair (u, v) lies in the interpretation of R in $\mathbf{A}$ will be indifferently denoted by $(u, v) \in R(\mathbf{A})$ or $\mathbf{A} \models R(u, v)$. More generally, for a formula $\varphi(x_1, \dots, x_k)$, a Σ-structure $\mathbf{A}$, an integer $\ell < k$ and $a_1, \dots, a_\ell \in A$ we define

$$\varphi(\mathbf{A}, a_1, \dots, a_\ell) := \{(x_1, \dots, x_{k-\ell}) \in A^{k-\ell} : \mathbf{A} \models \varphi(x_1, \dots, x_{k-\ell}, a_1, \dots, a_\ell)\}.$$

Let $\mathbf{A} = (A, R_1(\mathbf{A}), \dots, R_s(\mathbf{A}))$ be a Σ-structure and let $X \subseteq A$. The *substructure* of $\mathbf{A}$ *induced by* X is the Σ-structure $\mathbf{A}[X] = (X, R_1(\mathbf{A}) \cap X^{r_1}, \dots, R_k(\mathbf{A}) \cap X^{r_s})$.

Graphs are structures with a single binary relation E encoding adjacency; this relation is anti-reflexive and symmetric. Graphs of particular interest in this paper are rooted trees. For a rooted tree Y, we denote by $I(Y)$ the set of internal nodes of Y, by $L(Y)$, the set of leaves of Y, by $r(Y)$, the root of Y, and by $\preceq_Y$, the partial order on $V(Y)$ defined by $u \preceq_Y v$ if the unique path in Y linking $r(Y)$ and v contains u (i.e., if $u = v$ or u is an *ancestor* of v in Y). For a non-root vertex v, we further denote by $\pi_Y(v)$ the *parent* of v, which is the unique neighbor of v smaller than v with respect to $\preceq_Y$. A rooted *binary* tree is a rooted tree such that every internal node has exactly two children.

Partial orders are structures with a single anti-symmetric and transitive binary relation $\prec$. Particular partial orders will be of interest here. *Linear orders* are partial orders such that $\forall x \forall y\ (x \prec y \vee y \prec x \vee y = x)$. *Tree orders* are partial orders that satisfy the following axioms: $\forall x \forall y \forall z\ \big((x \prec z \wedge y \prec z) \to (x \prec y \vee y \prec x \vee x = y)\big)$ and $\exists r \forall x\ (x = r \vee r \prec x)$. It will be convenient to use $\preceq, \succ, \succeq$ with their obvious meaning. Let $(X, \prec)$ be a tree-order.

The *infimum* $\inf(u,v)$ of two elements $u,v \in X$ is the unique element $w \in X$ such that $w \preceq u, w \preceq v$, and $\forall z\ \big((z \preceq u \wedge z \preceq v) \to z \preceq w\big)$. Note that $\inf(x,y)$ is first-order definable from $\prec$, hence can be used in our formulas. A *binary tree order* is a tree-order $(X,\prec)$ that satisfies the axiom $\forall x\, \exists y_1 \exists y_2\ \big(y_1 \neq y_2 \wedge \forall z\ (z \succ x) \to (z \succeq y_1 \vee z \succeq y_2)\big)$.

Ordered graphs are structures with two binary relations, E and $<$, where E defines a graph and $<$ defines a linear order. We denote ordered graphs as $G^< = (V,E,<)$.

A *permutation* is represented as a structure $\sigma = (V, <_1, <_2)$, where V is a finite set and where $<_1$ and $<_2$ are two linear orders on this set (see e.g. [7, 1]). Two permutations $\sigma = (V, <_1, <_2)$ and $\sigma' = (V', <_1', <_2')$ are *isomorphic* if there is a bijection between V and V' preserving both linear orders. Let $X \subseteq V$. The *sub-permutation* of σ induced by X is the permutation on X defined by the two linear orders of σ restricted to X. The isomorphism types of the sub-permutations of a permutation σ are the *patterns* of σ. A class $\mathscr{P}$ of (isomorphism types of) permutations is *hereditary* (or *closed*) if it is closed under taking sub-permutations. A *permutation class* is a hereditary class of permutations. A permutation class is *proper* if it is not the class of all permutations. Note that the terms "class of permutations" and "permutation class" are not equivalent, the second referring to a hereditary class of permutations, as it is customary.

2.2 Interpretations

Let Σ, Σ' be signatures. A (simple) *interpretation* I of Σ'-structures in Σ-structures is defined by a formula $\rho_0(x)$, and a formula $\rho_{R'}(x_1,\dots,x_k)$ for each k-ary relation symbol $R' \in \Sigma'$ (where, by formula, we mean a first-order formula in the language of Σ-structures). Let I be an interpretation of Σ'-structures in Σ-structure, where $\Sigma' = \{R_1', \dots, R_s'\}$. For each Σ-structure $\mathbf{A}$ we denote by $\mathsf{I}(\mathbf{A}) = (\rho_0(\mathbf{A}), \rho_{R_1'}(\mathbf{A}), \dots, \rho_{R_s'}(\mathbf{A}))$ the Σ'-structure interpreted by I in $\mathbf{A}$. Similarly, for a class $\mathscr{C}$ we denote by $\mathsf{I}(\mathscr{C})$ the set $\{\mathsf{I}(\mathbf{A}) : \mathbf{A} \in \mathscr{C}\}$.

We denote by $\mathsf{Reduct}_{\Sigma^+ \to \Sigma}$ (or simply Reduct when Σ and Σ^+ are clear from context) the interpretation that "forgets" the relations in $\Sigma^+ \setminus \Sigma$ while preserving all the other relations and the domain. For a Σ^+-structure $\mathbf{B}$, the Σ-structure $\mathsf{Reduct}(\mathbf{B})$ is called the Σ-*reduct* (or simply *reduct* if Σ is clear from the context) of $\mathbf{B}$. A class $\mathscr{C}$ is a *reduct* of a class $\mathscr{D}$ if $\mathscr{C} = \mathsf{Reduct}(\mathscr{D})$. Conversely, a class $\mathscr{D}$ is an *expansion* of $\mathscr{C}$ if $\mathscr{C}$ is a reduct of $\mathscr{D}$.

Another important interpretation is $\mathsf{Gaifman}_\Sigma$ (or simply $\mathsf{Gaifman}$ when Σ is clear from context), which maps a Σ-structure $\mathbf{A}$ to its *Gaifman graph*, whose vertex set is A and whose edge set is the set of all pairs of vertices included in a tuple of some relation.

Note that an interpretation of Σ_2-structures in Σ_1-structure naturally defines an interpretation of Σ_2^+-structures in Σ_1^+-structures if $\Sigma_2^+ \setminus \Sigma_2 = \Sigma_1^+ \setminus \Sigma_1$ by leaving the relations in $\Sigma_1^+ \setminus \Sigma_1$ unchanged (that is, by considering $\rho_R(x_1,\dots,x_k) = R(x_1,\dots,x_k)$ for these relations).

2.3 Transductions

Let Σ, Σ' be signatures. A *simple transduction* T from Σ-structures to Σ'-structures is defined by a simple interpretation I_T of Σ'-structures in Σ^+-structures, where Σ^+ is a signature obtained from Σ by adding finitely many marks. For a Σ-structure $\mathbf{A}$, we denote by $\mathsf{T}(\mathbf{A})$ the set of all $\mathsf{I}_\mathsf{T}(\mathbf{B})$ where $\mathbf{B}$ is a Σ^+-structure with reduct $\mathbf{A}$: $\mathsf{T}(\mathbf{A}) = \{\mathsf{I}_\mathsf{T}(\mathbf{B}) : \mathsf{Reduct}(\mathbf{B}) = \mathbf{A}\}$. Let $k \in \mathbb{N}$. A k-*blowing* of a Σ-structure $\mathbf{A}$ is the Σ'-structure $\mathbf{B} = \mathbf{A} \bullet k$, where Σ' is the signature obtained from Σ by adding a new binary relation $\sim$ encoding an equivalence relation. The domain of $\mathbf{A} \bullet k$ is $B = A \times [k]$, and, denoting p the projection $A \times [k] \to A$ we have, for all $x, y \in B$, $\mathbf{B} \models x \sim y$ if $p(x) = p(y)$, and (for $R \in \Sigma$) $\mathbf{B} \models R(x_1,\dots,x_k)$ if

$\mathbf{A} \models R(p(x_1), \ldots, p(x_k))$. A *copying transduction* is the composition of a k-blowing and a simple transduction; the integer k is the *blowing factor* of the copying transduction T and is denoted by $\mathrm{bf}(\mathsf{T})$. It is easily checked that the composition of two copying transductions is again a copying transduction. In the following by the term transduction we mean a copying transduction. Note that for every transduction T from Σ-structure to Σ', for every Σ-structure $\mathbf{A}$ and for every Σ'-structure $\mathbf{B} \in \mathsf{T}(\mathbf{A})$ we have $|B| \leq \mathrm{bf}(\mathsf{T})\,|A|$.

Let T, T' be transductions from Σ-structures to Σ'-structures, and let $\mathscr{C}$ be a class of Σ-structures. The transduction T' *subsumes* the transduction T *on* $\mathscr{C}$ if $\mathsf{T}'(\mathbf{A}) \supseteq \mathsf{T}(\mathbf{A})$ for all $\mathbf{A} \in \mathscr{C}$. If $\mathscr{C}$ is a class of Σ-structures we define $\mathsf{T}(\mathscr{C}) = \bigcup_{\mathbf{A} \in \mathscr{C}} \mathsf{T}(\mathbf{A})$. We say that a class $\mathscr{D}$ of Σ'-structure is a T-*transduction* of $\mathscr{C}$ if $\mathscr{D} \subseteq \mathsf{T}(\mathscr{C})$ and, more generally, the class $\mathscr{D}$ is a *transduction* of the class $\mathscr{C}$, and we write $\mathscr{C} \longrightarrow\!\!\!\!\!\rightarrow \mathscr{D}$, if there exists a transduction T such that $\mathscr{D}$ is a T-transduction of $\mathscr{C}$. Note that we only require the inclusion of $\mathscr{D}$ in $\mathsf{T}(\mathscr{C})$. The class $\mathscr{D}$ is a *c-bounded* T-transduction of the class $\mathscr{C}$ if, for every $\mathbf{B} \in \mathscr{D}$ there exists $\mathbf{A} \in \mathscr{C}$ with $\mathbf{B} \in \mathsf{T}(\mathbf{A})$ and $|A| \leq c\,|B|$. Two classes $\mathscr{C}$ and $\mathscr{D}$ are *transduction equivalent* if each is a transduction of the other. A *transduction pairing* of two classes $\mathscr{C}$ and $\mathscr{D}$ is a pair (D, C) of (copying) transductions, such that

$$\forall \mathbf{A} \in \mathscr{C} \quad \exists \mathbf{B} \in \mathsf{D}(\mathbf{A}) \cap \mathscr{D} \quad \mathbf{A} \in \mathsf{C}(\mathbf{B}) \quad \text{and} \quad \forall \mathbf{B} \in \mathscr{D} \quad \exists \mathbf{A} \in \mathsf{C}(\mathbf{B}) \cap \mathscr{C} \quad \mathbf{B} \in \mathsf{D}(\mathbf{A}).$$

We denote by $\mathscr{C} \rightleftharpoons \mathscr{D}$ the existence of a transduction pairing of $\mathscr{C}$ and $\mathscr{D}$. Remark that if (C, D) is a transduction pairing then C is $\mathrm{bf}(\mathsf{D})$-bounded and D is $\mathrm{bf}(\mathsf{C})$-bounded.

2.4 Twin-width

Inspired by a width invariant defined on permutations by Guillemot and Marx [11], the twin-width invariant tww has been recently introduced by Bonnet, Kim, Thomassé, and Watrigant [6]. Graph twin-width is originally defined using a sequence of near-twin identifications. The twin-width of the graph intuitively measures the accumulated errors (kept track of by so-called *red edges*) made by the identifications. Our definition for binary relational structures will be equivalent when restricted to undirected graphs, but will slightly differ from the definition of the twin-width of binary structures (via matrix encodings) given in [6], though linearly tied. Note that forbidding unary relations does not harm in our context, as transductions can freely reintroduce any number of unary relations.

In order to define twin-width, we first need to introduce some preliminary notions, which generalize the notion of trigraphs (i.e., graphs with some red edges) introduced in [6]. Let Σ be a binary relational signature. The signature Σ^* is obtained by adding, for each binary relation symbol R a new binary relation symbol R^*. The symbol R^* will always be interpreted as a symmetric relation and plays for R the role of red edges in [6].

Let $\mathbf{A}$ be a Σ^*-structure and let u and v be vertices of $\mathbf{A}$. The vertices u, v are *R-clones* for a vertex w and a relation $R \in \Sigma$ if we have $\mathbf{A} \models \big(R(u,w) \leftrightarrow R(v,w)\big) \wedge \big(R(w,u) \leftrightarrow R(w,v)\big)$ and no pair in R^* contains both w and either u or v. The Σ^*-structure $\mathbf{A}'$ obtained by *contracting* u and v into a new vertex z is defined as follows:

- $A' = A \setminus \{u, v\} \cup \{z\}$;
- $R(\mathbf{A}') \cap (A' \setminus \{z\}) \times (A' \setminus \{z\}) = R(\mathbf{A}) \cap (A \setminus \{u, v\}) \times (A \setminus \{u, v\})$ for all $R \in \Sigma^*$;
- for every vertex $w \in A' \setminus \{z\}$ and every $R \in \Sigma$ such that u and v are R-clones for w, we let $\mathbf{A}' \models R(w, z)$ if $\mathbf{A} \models R(u, z)$ and $\mathbf{A}' \models R(z, w)$ if $\mathbf{A} \models R(z, u)$;
- otherwise, for every vertex $w \in A' \setminus \{z\}$ and every $R \in \Sigma$ such that u and v are not R-clones for w we let $\mathbf{A}' \models R^*(w, z) \wedge R^*(z, w)$.

A *d-sequence* for a Σ-structure $\mathbf{A}$ is a sequence $\mathbf{A}_n, \ldots, \mathbf{A}_1$ of Σ^*-structures such that: $\mathbf{A}_n$ is isomorphic to $\mathbf{A}$; $\mathbf{A}_1$ is the Σ^*-structure with a single element; for every $1 \leq i < n$, $\mathbf{A}_i$ is obtained from $\mathbf{A}_{i+1}$ by performing a single contraction; for every $1 \leq i < n$ and every $v \in A_i$, the sum of the degrees in relations $R^* \in \Sigma^* \setminus \Sigma$ of v in $\mathbf{A}_i$ is less or equal to d (the degree of v in relation R^* is defined as the degree of the undirected graph $(A, R^*(\mathbf{A}))$. The minimum d such that there exists a d-sequence for a Σ-structure $\mathbf{A}$ is the *twin-width* $\mathrm{tww}(\mathbf{A})$ of $\mathbf{A}$. A crucial property of twin-width is the following result.

▶ **Theorem 2.1** ([6]). *Let $\mathscr{C}, \mathscr{D}$ be classes of binary structures. If $\mathscr{C}$ has bounded twin-width and $\mathscr{D}$ is a transduction of $\mathscr{C}$, then $\mathscr{D}$ has bounded twin-width.*

3 Transductions of sparse classes

Sparse classes of graphs have many nice properties. The next property will be of particular interest here. Recall that a *star coloring* of a graph G is a proper coloring of G such that any two color classes induce a star forest (i.e., a disjoint union of stars); the *star chromatic number* $\chi_{\mathrm{st}}(G)$ of G is the minimum number of colors in a star coloring of G. Note that a star coloring of a graph with c defines a partition of the edge set into $\binom{c}{2}$ star forests. Although we are interested only in binary relational structures in this paper, the next lemma holds (and is proved) for general relational signatures.

▶ **Lemma 3.1.** *Let Σ be a relational signature, let $\mathscr{C}$ be a class of Σ-structures, and let c be an integer. There exists a simple transduction* Unfold_c *from graphs to Σ-structures such that if graphs in $\mathscr{C}$ have star chromatic number at most c, then* $(\mathsf{Unfold}_c, \mathsf{Gaifman})$ *is a transduction pairing of* $(\mathscr{C}, \mathsf{Gaifman}(\mathscr{C}))$.

Proof. Let $c = \sup\{\chi_{\mathrm{st}}(G) : G \in \mathsf{Gaifman}(\mathscr{C})\} < \infty$. Let $\mathbf{A} \in \mathscr{C}$, let $G = \mathsf{Gaifman}(\mathbf{A})$, and let $\gamma : V(G) \to [c]$ be a star coloring of G. In G, any two color classes induce a star forest, which we orient away from their centers. This way we get an orientation $\vec{G}$ of G such that for every vertex v and every in-neighbor u of v, the vertex u is the only in-neighbor of v with color $\gamma(u)$. Let $R \in \Sigma$ be a relation of arity k. For each $(u_1, \ldots, u_k) \in R(\mathbf{A})$, $u_1, \ldots, u_k$ induce a tournament in $\vec{G}$. Every tournament has at least one directed Hamiltonian path [16]. We fix one such Hamiltonian path and let $p(u_1, \ldots, u_k)$ be the index of the last vertex in the path. Let $a = p(u_1, \ldots, u_k)$, let $(c_1, \ldots, c_k) = (\gamma(u_1), \ldots, \gamma(u_k))$. Then there exists exactly one clique of size k containing u_a with vertices colored $c_1, \ldots, c_k$. Indeed, let $(u_{i_1}, \ldots, u_{i_k})$ be the Hamiltonian path associated to the tournament induced by $u_1, \ldots, u_k$ in $\vec{G}$. Then $u_{i_k} = u_a$ and, for each $1 \leq j < k$ the vertex u_{i_j} is the unique neighbor of $u_{i_{j+1}}$ with color c_{i_j}. For each relation $R \in \Sigma$ with arity k and each $(u_1, \ldots, u_k) \in R(\mathbf{A})$ we put at $v = u_{p(u_1, \ldots, u_k)}$ a mark $M^R_{\gamma(u_1), \ldots, \gamma(u_k)}$. We further put at each vertex v a mark $C_{\gamma(v)}$. Then the structure $\mathbf{A}$ is reconstructed by the transduction Unfold_c defined by the formulas

$$\rho_R(x_1, \ldots, x_k) := \bigvee_{c_1, \ldots, c_k} \Big(\bigwedge_{1 \leq j \leq k} C_{c_i}(x_i) \wedge \bigwedge_{1 \leq i < j \leq k} E(x_i, x_j) \wedge \bigvee_{1 \leq i \leq k} M^R_{c_1, \ldots, c_k}(x_i) \Big).$$

◀

Note that if the condition of Lemma 3.1 is tight in the sense that if a class $\mathscr{C}$ of undirected graphs has high girth (meaning that the girth grows with the order with the graph) and unbounded star chromatic number (hence unbounded oriented chromatic number) then the class $\vec{\mathscr{C}}$ of all orientations of the graphs in $\mathscr{C}$ is not a transduction of $\mathscr{C}$ [15].

Lemma 3.1 will be particularly significant in conjunction with the following results. (We refrain here from formally defining classes with bounded expansion, as we will use only the properties of these classes expressed by the next two theorems.)

▶ **Theorem 3.2** ([4]). *Every degenerate class with bounded twin-width has bounded expansion.*

▶ **Theorem 3.3** ([14]). *Every class with bounded expansion has bounded star chromatic number.*

4 Twin-models

In this section we formalize the notions of twin-models and ranked twin-models, which are reminiscent of the "ordered union trees" and "interval biclique partitions" adopted in [3]. This structure will allow to encode a contraction sequence and to give an alternative definition of twin-width.

4.1 Twin-models, ranking, layers, and width

▶ **Definition 4.1** (twin-model). *Let $\Sigma = (R_1, \dots, R_k)$ be a binary relational signature. A Σ-*twin-model *(or simply a* twin-model *when Σ is clear from the context) is a tuple $(Y, Z_{R_1}, \dots, Z_{R_k})$ where Y is a rooted binary tree and each Z_{R_i} is a binary relation satisfying the following minimality and consistency conditions:*

- (minimality) *if $(u, v) \in Z_{R_i}$, then there exists no $(u', v') \neq (u, v)$ with $u' \preceq_Y u$, $v' \preceq_Y v$ and $(u', v') \in Z_{R_i}$;*
- (consistency) *if a traversal of a cycle γ in $Y \cup \bigcup_i Z_{R_i}$ traverses all the Y-edges (of γ) away from the root, then γ contains two consecutive edges in $\bigcup_i Z_{R_i}$.*

A twin-model $(Y, Z_{R_1}, \dots, Z_{R_k})$ defines *the Σ-structure $\mathbf{A}$ (or $(Y, Z_{R_1}, \dots, Z_{R_k})$ is a* twin-model of $\mathbf{A}$*) if $A = L(Y)$ and, for each $R_i \in \Sigma$, $R_i(\mathbf{A})$ is the set of all pairs (u, v) such that there exists $u' \preceq_Y u$ and $v' \preceq_Y v$ with $(u', v') \in Z_{R_i}$.*

▶ **Definition 4.2** (ranking, boundaries, and layers). *Let $(Y, Z_{R_1}, \dots, Z_{R_k})$ be a twin-model of a Σ-structure $\mathbf{A}$ with $|A| = n$. A* ranking *τ of the twin-model $(Y, Z_{R_1}, \dots, Z_{R_k})$ is a mapping from $V(Y)$ to $[n]$ that satisfies the following labeling, monotonicity, and synchronicity conditions:*

- (labeling) *the function τ restricted to $I(Y)$ is a bijection with $[n-1]$, and is equal to n on $L(Y)$;*
- (monotonicity) *If $u \prec_Y v$, then $\tau(u) < \tau(v)$;*
- (synchronicity) *If $(u, v) \in Z_{R_i}$, then $\max(\tau(\pi_Y(u)), \tau(\pi_Y(v))) < \min(\tau(u), \tau(v))$.*

A ranked twin-model *is a tuple $\mathfrak{T} = (Y, Z_{R_1}, \dots, Z_{R_k}, \tau)$, where $(Y, Z_{R_1}, \dots, Z_{R_k})$ is a twin-model, and τ is a ranking of $(Y, Z_{R_1}, \dots, Z_{R_k})$.*

For $1 < t \leq n$, the boundary *$\partial_t Y$ is the set $\partial_t Y = \{u \in V(Y) \mid \tau(u) \geq t \wedge \tau(\pi_Y(u)) < t\}$ and the* layer *$\mathbf{L}_t$ is the Σ^*-structure with vertex set $\partial_t Y$ and relations defined by*

$$R_i(\mathbf{L}_t) = \{(u, v) \in \partial_t Y \times \partial_t Y \mid \exists u' \preceq_Y u,\ \exists v' \preceq_Y v,\ (u', v') \in Z_{R_i}\}$$
$$R_i^*(\mathbf{L}_t) = \{(u, v) \in \partial_t Y \times \partial_t Y \mid \exists u' \succeq_Y u,\ \exists v' \succeq_Y v,\ (u', v') \neq (u, v)$$
$$\text{and } \{(u', v'), (v', u')\} \cap Z_{R_i} \neq \emptyset\}.$$

For $t = 1$ we define the boundary $\partial_1 Y = \{r(Y)\}$ and the layer $\mathbf{L}_1$ as the Σ^*-structure with unique vertex $r(Y)$.

▶ **Definition 4.3** (width). *The* width *of a ranked twin-model* $\mathfrak{T} = (Y, Z_{R_1}, \dots, Z_{R_k}, \tau)$ *is defined as*

$$\text{width}(\mathfrak{T}) = \max_{t \in [n]} \max_{v \in L_t} \sum_{R_i \in \Sigma} |R_i^*(\mathbf{L}_t, v)|.$$

At first sight the consistency condition of a twin-model (of $\mathbf{A}$) may seem contrived. One may for instance wonder if the minimality and consistency conditions are not simply equivalent to the property that every $(u, v) \in R_i(\mathbf{A})$ is *realized* by a unique unordered pair u', v' with $(u', v') \in Z_{R_i}$, $u' \preceq_Y u$, and $v' \preceq_Y v$. In case the structure $\mathbf{A}$ encodes a simple undirected graph G (with signature $\Sigma = (E)$), we would simply impose that the edges of Z_E partition the edges of G into bicliques.

In Figure 2 we give a small example that shows that this property is not strong enough to always yield a ranking. This illustrates why the consistency condition is what we want (no more, no less) and also serves as a visual support for the notions of contraction sequence, twin-model, and ranking.

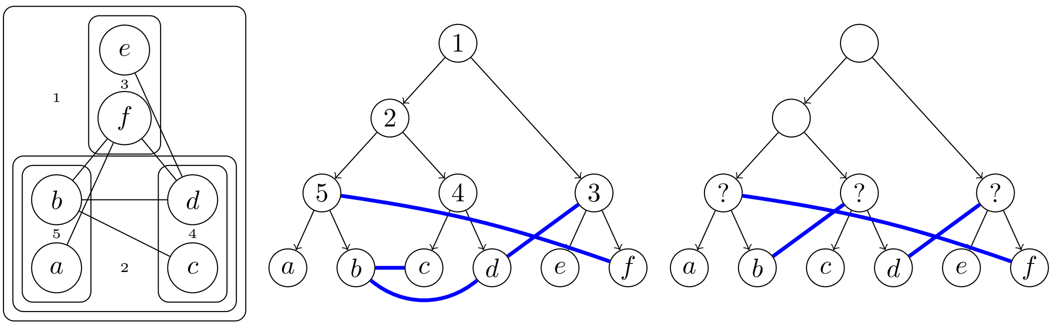

Figure 2 Left: A 6-vertex graph and a contraction sequence, where the tiny digit in each box indicates the number of remaining vertices when the vertex appears. Center: A twin-model of the graph, where the edges of Z_E are in bold blue, and a ranking (for the internal nodes) of this twin-model that actually matches the contraction sequence. Right: A *flawed twin-model* where the edge set E is indeed partitioned by the pairs of Z_E. Here no ranking is possible: By the symmetry, one just needs to consider the labeling of the parent of a, b with 5, and that then the edges bc and bd cannot be realized. There is indeed a cycle $b?d?f?$ with all the tree arcs oriented the same way, and without two consecutive edges of Z_E. On the contrary all such cycles in the central tree have two consecutive edges of Z_E, like $5bd3f$ has $(b, d), (d, 3) \in Z_E$.

4.2 From a contraction sequence to a twin-model

In this section we prove that every d-contraction sequence of a Σ-structure $\mathbf{A}$ defines a ranked twin-model of $\mathbf{A}$ with width at most d.

A d-sequence $\mathbf{A}_n, \dots, \mathbf{A}_1$ for a Σ-structure $\mathbf{A}$ defines a rooted binary tree Y with vertex set $V(Y) = \bigcup_i A_i$ and set of leaves $L(Y) = A_n$ as follows: for each $i \in [n-1]$ let z_i be the vertex of A_i and u_i, v_i be the vertices of A_{i+1} such that z_i results from the contraction of u_i and v_i in $\mathbf{A}_{i+1}$. Then $I(Y) = \{z_i : i \in [n-1]\}$, $r(Y) = z_1$, and the children of z_i in Y are the vertices u_i and v_i.

For each relation $R \in \Sigma$ we define a binary relation Z_R on $V(Y)$ as follows. Let z_i be the vertex of A_i resulting from the contraction of u_i and v_i in A_{i+1}. If $(u_i, v_i) \in R(\mathbf{A}_{i+1})$, then $(u_i, v_i) \in Z_R$. If u_i and v_i are not R-clones for w, then the pairs involving w and u_i or v_i in $R(\mathbf{A}_{i+1})$ are copied in Z_R. Intuitively, Z_R collects the R-relations when they just appear (in the order $\mathbf{A}_1, \dots, \mathbf{A}_n$). We further define $Z = \bigcup_{R\in\Sigma} Z_R$ and the function $\tau\colon V(Y) \to [n]$ by $\tau(v) = n$ if $v \in L(Y)$ and $\tau(z_i) = i$. Note that for each $i \in [n]$ and non-root vertex v of Y, we have $v \in A_i$ if and only if $\tau(\pi_Y(v)) < i \leq \tau(v)$.

▶ **Lemma 4.4.** *Every d-sequence $\mathbf{A}_n, \dots, \mathbf{A}_1$ defines a ranked twin-model with width at most d.*

Proof.

▷ Claim 4.5. The function τ satisfies the labeling, monotonicity, and synchronicity conditions.

Proof. The first two conditions are straightforward. Let $(u, v) \in Z_R$. Let $i \in [n-1]$ be such that (u, v) appears in $\mathbf{A}_i$ for the first time. As $u, v \in A_i$ we have both $\tau(\pi_Y(u)) < i \leq \tau(u)$ and $\tau(\pi_Y(v)) < i \leq \tau(v)$, i.e., the synchronicity condition holds. ◁

▷ Claim 4.6. The relations Z_R $(R \in \Sigma)$ satisfy the minimality and consistency conditions.

Proof. The minimality condition follows directly from the definition. Let $\vec{H}$ be the oriented graph obtained from Y by orienting all the edges from the root and adding, for each $R \in \Sigma$ and each pair $(u, v) \in Z_R$ the arcs $\pi_Y(u)v$ and $\pi_Y(v)u$ whenever they do not exist. It follows from the monotonicity and synchronicity conditions that $\vec{H}$ is acyclically oriented. Indeed, any arc (x, y) in $\vec{H}$ satisfies $\tau(x) < \tau(y)$.

Assume towards a contradiction that in $Y \cup \bigcup_{R\in\Sigma} Z_R$ one can find a cycle γ such that the orientation of the Y-edges is consistent with a traversal of γ and γ does not contain two consecutive edges in $\bigcup_{R\in\Sigma} Z_R$. By replacing in γ each group formed by an edge in $\bigcup_{R\in\Sigma} Z_R$ and its preceding edge in Y by the corresponding arc in $\vec{H}$ we obtain a circuit in $\vec{H}$, contradicting its acyclicity. Hence the relations Z_R satisfy the consistency condition. ◁

The next claim is immediate from the definition and ends the proof of the lemma.

▷ Claim 4.7. The ranked twin-model $(Y, Z_{R_1}, \dots, Z_{R_k}, \tau)$ derived from a d-sequence $\mathbf{A}_n, \dots, \mathbf{A}_1$ has width at most d.

◀

4.3 Properties of twin-models

In this section we establish two properties of twin-models. The first one is the equality of the minimum width of a twin-model with the twin-width of a structure; the second one is that twin-models of structures with bounded twin-width are degenerate.

▶ **Lemma 4.8.** *Every twin-model has a ranking, and the twin-width of a Σ-structure $\mathbf{A}$ is the minimum width of a ranked twin-model of $\mathbf{A}$.*

Proof. The following claim, which asserts that no Z_{R_i} "crosses" the boundaries, will be quite helpful.

▷ Claim 4.9. Let $t \in [n-1]$ and let $u, v \in \partial_t Y$. Then there exists no pair $(u', v') \in Z_{R_i}$ with $u' \prec_Y u$ and $v' \succ_Y v$.

Proof. Assume $(u', v') \in Z_{R_i}$ and $v' \succ_Y v$. By the synchronicity property we have $\tau(u') > \tau(\pi_Y(v')) \geq \tau(v) \geq t$, contradicting $\tau(u') \leq \tau(\pi_Y(u)) < t$. ◁

For a Σ^*-structure $\mathbf{A}$ and $R \in \Sigma$ we define

$$\overline{R}(\mathbf{A}) = \{(u, v) \in A^2 : \{(u, v), (v, u)\} \cap (R(\mathbf{A}) \cup R^*(\mathbf{A})) \neq \emptyset\}.$$

▷ Claim 4.10. Let $\mathbf{L}_1, \dots, \mathbf{L}_n$ be the layers of a ranked tree model $\mathfrak{T}$ of a Σ-structure $\mathbf{A}$. Then there exists a contraction sequence $\mathbf{A}_n, \dots, \mathbf{A}_1$ of $\mathbf{A}$ with $A_i = L_i$ and, for each $R \in \Sigma$, $\overline{R}(\mathbf{A}_i) = \overline{R}(\mathbf{L}_i)$, $R(\mathbf{L}_i) \subseteq R(\mathbf{A}_i)$, and $R^*(\mathbf{L}_i) \supseteq R^*(\mathbf{A}_i)$.

Proof. For $i \in [n-1]$, the Σ-structure $\mathbf{A}_i$ is obtained from $\mathbf{A}_{i+1}$ by contracting the pair of vertices u_i, v_i into w_i, where w_i is the vertex of Y with $\tau(w_i) = i$ and u_i and v_i are the two children of w_i in Y. It is easily checked that $A_i = L_i$. Let z be a vertex of A_i different from w_i. Then $(z, w_i) \in \overline{R}(\mathbf{A}_i)$ if there exists a leaf $w' \succeq_Y w_i$ and a leaf $z' \succeq_Y z$ such that $\{(w', z'), (z', w')\} \cap R(\mathbf{A}) \neq \emptyset$. As $\mathfrak{T}$ is a twin-model of $\mathbf{A}$ this means that there exists $w'' \preceq_Y w'$ and $z'' \preceq_Y z'$ with $(w'', z'') \in Z_R$ or $(z'', w'') \in Z_R$. As $\preceq_Y$ is a tree-order, w_i and w'' are comparable, as well as z and z''. From this and Claim 4.9 it follows $\{(z, w_i), (w_i, z)\} \subseteq \overline{R}(\mathbf{L}_i)$.

We now prove $R(\mathbf{L}_i) \subseteq R(\mathbf{A}_i)$ by reverse induction on i. For $i = n$ we have $R(\mathbf{L}_i) = R(\mathbf{A}_i) = R(\mathbf{A})$. Let $i \in [n-1]$ and let u_i, v_i, w_i be defined as above. If $(w_i, z) \in R(\mathbf{L}_i)$, then there exists $w' \preceq_Y w_i$ and $z' \preceq_Y z$ with $(w', z') \in Z_R$ thus we have also $(u_i, z) \in R(\mathbf{L}_{i+1})$ and $(v_i, z) \in R(\mathbf{L}_{i+1})$. By induction we deduce $(u_i, z) \in R(\mathbf{A}_{i+1})$ and $(v_i, z) \in R(\mathbf{A}_{i+1})$. Similarly, if $(z, w_i) \in R(\mathbf{L}_i)$, then $(z, u_i) \in R(\mathbf{A}_{i+1})$ and $(z, v_i) \in R(\mathbf{A}_{i+1})$. Thus u_i and v_i are R-clones for z hence if $(w_i, z) \in R(\mathbf{L}_i)$, then $(w_i, z) \in R(\mathbf{A}_i)$ and if $(z, w_i) \in R(\mathbf{L}_i)$, then $(z, w_i) \in R(\mathbf{A}_i)$. It follows that we have $R(\mathbf{L}_i) \subseteq R(\mathbf{A}_i)$. Thus we have

$$\begin{aligned} R^*(\mathbf{L}_i) &= \overline{R}(\mathbf{L}_i) \setminus \{(u, v) : \{(u, v), (v, u)\} \cap R(\mathbf{L}_i) = \emptyset\} \\ &\supseteq \{(u, v) : \overline{R}(\mathbf{A}_i) \setminus \{(u, v) : \{(u, v), (v, u)\} \cap R(\mathbf{A}_i) = \emptyset\} = R^*(\mathbf{A}_i). \end{aligned}$$ ◀

We are now able to prove the first part of the statement.

▷ Claim 4.11. Every twin-model has a ranking.

Proof. Consider the oriented graph $\vec{H}$ obtained from orienting Y from the root and adding, for each $R \in \Sigma$ and each pair $(u, v) \in Z_R$, an arc $\pi(u)v$ and an arc $\pi(v)u$ (whenever they do not exist). Assume for contradiction that $\vec{H}$ contains a directed circuit. Replace each arc of the form $\pi(u)v$ of this circuit (with $(u, v) \in Z_R$) by the path $(\pi(u)u, uv)$ in the twin-model. This way we obtain a closed walk in $Y \cup \bigcup_{R\in\Sigma} Z_R$ traversing all edges of Y away from the root and no two consecutive edges are in $\bigcup_{R\in\Sigma} Z_R$, contradicting the consistency assumption. Thus $\vec{H}$ is acyclic and a topological ordering of $\vec{H}[I(Y)]$ extends to a labeling $\tau \colon V(H) \to [n]$ that is bijective between $I(Y)$ and $[n-1]$, equal to n on $L(Y)$, and increasing with respect to every arc of $\vec{H}$. This directly implies both the monotonicity and the synchronicity properties. ◁

We are now able to complete the proof of the lemma. According to Lemma 4.4, every d-sequence for $\mathbf{A}$ defines a ranked twin-model with width at most d. Conversely, every ranked twin-model for $\mathbf{A}$ with width d' defines a sequence of layers $\mathbf{L}_t$ with $\max_{v\in L_t} \sum_{R_i\in\Sigma} |R_i^*(\mathbf{L}_t, v)| \leq d'$ and, by Claim 4.10, a d'-sequence for $\mathbf{A}$. ◀

The following easy remark will be useful.

▷ Claim 4.12. Let $\mathfrak{Y} = (Y, Z_{R_1}, \dots, Z_{R_k}, \tau)$ be a ranked twin-model of a Σ-structure $\mathbf{A}$ and let $X \subseteq A$. Let Y' be the subtree of Y induced by all the vertices in X and their ancestors in Y, let Z'_{R_i} be the subset of all pairs in $Z_{R_i} \cap (Y' \times Y')$, and let τ' be the mapping from Y' to $[|X|]$ such that for every $x, y \in V(Y')$ we have $\tau(x) < \tau(y) \iff \tau'(x) < \tau'(y)$. Then $\mathfrak{Y}' = (Y', Z'_{R_1}, \dots, Z'_{R_k}, \tau')$ is a ranked twin-model of $\mathbf{A}[X]$, whose width is not larger than the one of $\mathfrak{Y}$.

▶ **Lemma 4.13.** *The Gaifman graph of a twin-model of a Σ-structure with width d is $d+k+1$-degenerate, where $k = |\Sigma|$.*

Proof. Let $\mathbf{A} = (A, R_1(\mathbf{A}), \dots, R_k(\mathbf{A}))$ be a Σ-structure, let $\mathfrak{T} = (Y, Z_{R_1}, \dots, Z_{R_k}, \tau)$ be a ranked twin-model of $\mathbf{A}$ with width d, and let G be the Gaifman graph of $(Y, Z_{R_1}, \dots, Z_{R_k})$. The ranked twin-model $\mathfrak{T}$ (with layers $\mathbf{L}_i$) defines a d-sequence $\mathbf{A}_n, \dots, \mathbf{A}_1$, where $A_i = L_i$ (see Lemma 4.8). Let z be the node with $\tau(z) = n-1$ and let u and v be its children. Each pair in Z_{R_i} containing u (except pairs containing to both u and v) gives rise (in $\mathbf{A}_{n-1}$) to an R_i^*-edge incident to z when contracting u and v. Thus the degree of u in G is at most $d+k+1$ (d for the sum of the degrees in the relations R_i^*, k for the pairs $(u,v) \in Z_{R_i}$, and 1 for the tree edge (u,z)). Then, in $G-u$, the vertex v has also degree at most $d+k+1$. Now we remark that removing u and v from Y, and redefining $\tau(x)$ as $\min(n-1, \tau(x))$, we get a ranked twin-model of $\mathbf{A}_{n-1}$ (minus R_i^*-edges) with width at most d, whose Gaifman graph is $G-u-v$. By induction, we deduce that G is $d+k+1$-degenerate. ◀

5 Full twin-models

To reconstruct a Σ-structure $\mathbf{A}$ from a twin-model $(Y, Z_{R_1}, \dots, Z_{R_k})$, we make use of the tree-order $\preceq_Y$ defined by Y. As this tree-order cannot be obtained as a first-order transduction of $(Y, Z_{R_1}, \dots, Z_{R_k})$ it will be convenient to introduce a variant of twin-models: the *full twin-model* associated to a twin-model $(Y, Z_{R_1}, \dots, Z_{R_k})$ is the structure $(V(Y), \prec_Y, Z_{R_1}, \dots, Z_{R_k})$. Let S be the simple interpretation of Σ-structures in full twin-models defined by formulas

$$\begin{aligned} \rho_0(x) &:= \neg(\exists y\ y \prec_Y x); \\ \rho_{R_i}(x,y) &:= \exists u\, \exists v\ (u \preceq_Y x) \wedge (v \preceq_Y y) \wedge Z_{R_i}(u,v). \end{aligned}$$

The following claim follows directly from the definition of a twin-model.

▷ Claim 5.1. If $\mathbf{T} = (X, \prec, Z_{R_1}, \dots, Z_{R_k})$ is a full twin-model of $\mathbf{A}$, then $\mathsf{S}(\mathbf{T}) = \mathbf{A}$ and $|A| = (|T|+1)/2$.

▶ **Lemma 5.2.** *Let $\mathfrak{T} = (Y, Z_{R_1}, \dots, Z_{R_k}, \tau)$ be a ranked twin-model, with associated full twin-model $\mathbf{T} = (V(Y), \prec_Y, Z_{R_1}, \dots, Z_{R_k})$. Then the width of $\mathbf{T}$ is at most twice the width of $(V(Y), \prec_Y, Z_{R_1}, \dots, Z_{R_k}, \tau)$.*

Proof. Let I_0, I_1 be copies of $I(Y)$ and let $p_i : I(Y) \to I_i$ be the "identity" for $i = 0, 1$. We define the binary rooted tree $\widehat{Y}$ with vertex set $V(\widehat{Y}) = V(Y) \cup I_1 \cup I_0$, leaf set $L(\widehat{Y}) = V(Y)$, root $r(\widehat{Y}) = p_o(r(Y))$, and parent function

$$\pi_{\widehat{Y}}(x) = \begin{cases} p_1 \circ \pi_Y(x) & \text{if } x \in L(Y) \\ p_0(x) & \text{if } x \in I(Y) \\ p_0 \circ p_1^{-1}(x) & \text{if } x \in I_1 \\ p_1 \circ \pi_Y \circ p_0^{-1}(x) & \text{if } x \in I_0 \setminus \{r(\widehat{Y})\} \\ x & \text{if } x = r(\widehat{Y}) \end{cases}$$

An informal description of $\widehat{Y}$ is that it is obtained by replacing every internal node v of Y by a *cherry* C_v (i.e., a complete binary tree on three vertices) such that one leaf of C_v remains a leaf in $\widehat{Y}$, the other leaf of C_v is linked to the "children of v," while the root of C_v is linked to the "parent of v" (provided v is not the root of Y).

We further define $\widehat{Z}_{\prec} = \{(v, p_1(v)) : v \in I(Y)\}$.

▷ Claim 5.3. $(\widehat{Y}, \widehat{Z}_{\prec}, Z_{R_1}, \dots, Z_{R_k})$ is a twin-model of $\mathbf{T}$.

Proof. We have $V(Y) = L(\widehat{Y})$. The minimality condition are obviously satisfied for $\widehat{Z}_{\prec}$ and Z_{R_i}. Let $\widehat{Z} = \widehat{Z}_{\prec} \cup \bigcup_i Z_{R_i}$. Consider a cycle $\widehat{\gamma}$ in $\widehat{Y} \cup \widehat{Z}$, with all the edges in $\widehat{Y}$ oriented away from the root. Assume for contradiction that no two edges in $\widehat{Z}$ are consecutive in $\widehat{\gamma}$. Then either $\widehat{\gamma}$ contains a directed path of $\widehat{Y}$ linking to vertices in $L(Y)$ or a directed path of $\widehat{Y}$ linking a vertex in $I(Y)$ to a distinct vertex in $V(Y)$. As no such directed paths exist in $\widehat{Y}$ we are led to a contradiction. Hence $(\widehat{Y}, \widehat{Z}_{\prec}, Z_{R_1}, \dots, Z_{R_k})$ satisfies the consistency condition.

For $u, v \in V(Y)$ there exists $u' \preceq_{\widehat{Y}} u$ and $v' \preceq_{\widehat{Y}} v$ with $(u', v') \in Z_{R_i}$ if and only if $(u, v) \in Z_{R_i}$. For $u, v \in V(Y)$ there exists $u' \preceq_{\widehat{Y}} u$ and $v' \preceq_{\widehat{Y}} v$ with $(u', v') \in \widehat{Z}_{\prec}$ if and only if $u' = u, v' = p_1(u)$, and $v' \preceq_{\widehat{Y}} v$, that is, if and only if $u \prec_Y v$. Hence $(\widehat{Y}, \widehat{Z}_{\prec}, Z_{R_1}, \dots, Z_{R_k})$ is a twin-model of $\mathbf{T}$. ◁

Let $n = |L(Y)|$. The next claim shows that we have much freedom in defining a ranking function for $(\widehat{Y}, \widehat{Z}_{\prec}, Z_{R_1}, \dots, Z_{R_k})$.

▷ Claim 5.4. If $\hat{\tau} : V(\widehat{Y}) \to [2n-1]$ satisfy the labeling and monotonicity conditions, then $\hat{\tau}$ is a ranking of $(\widehat{Y}, \widehat{Z}_{\prec}, Z_{R_1}, \dots, Z_{R_k})$.

Proof. Assume $(u, v) \in \widehat{Z}_{\prec}$. Then $\pi_{\widehat{Y}}(u) = \pi_{\widehat{Y}}(v)$ hence the synchronicity for $\widehat{Z}_{\prec}$ follows from monotonicity. Assume $(u, v) \in Z_{R_i}$. Then $\hat{\tau}(u) = \hat{\tau}(v) = 2n-1$ hence the synchronicity obviously holds. ◁

We now define $\hat{\tau} : V(\widehat{Y}) \to [2n-1]$ as follows: order the vertices $v \in I_1$ by increasing $\tau \circ p_1^{-1}(v)$. For each $v \in I_1$, insert the children of v in I_0 just after v, then add $r(\widehat{Y})$ in the very beginning. Numbering the vertices of $I_0 \cup I_1$ according this order defines $\hat{\tau}$ on $I(\widehat{Y})$. We extend this function to the whole $V(\widehat{Y})$ by defining $\hat{\tau}(v) = 2n-1$ for all $v \in L(\widehat{Y})$. By construction, the labeling and monotonicity properties hold hence $\hat{\tau}$ is a ranking of $(\widehat{Y}, \widehat{Z}_{\prec}, Z_{R_1}, \dots, Z_{R_k})$.

Consider a time $1 < \hat{t} < 2n-1$ and let v be the vertex with $\hat{\tau}(v) = \hat{t}$. If $v \in I_1$ we define $t = \tau \circ p_1^{-1}(v)$. Then $\partial_{\hat{t}}\widehat{Y} = p_1(\partial_t Y)$ and the degree for $Z^*_{R_i}$ in the layer of $\widehat{Y}$ at time $\hat{t}$ is at most the degree for $Z^*_{R_i}$ in the layer of Y at time t and the degree for $Z^*_{\prec}$ is null. If $v \in I_0$ we define $t = \tau \circ \pi_Y \circ p_0^{-1}(v)$. Then $\partial_{\hat{t}}\widehat{Y}$ is $p_1(\partial_t Y)$ in which we remove the parent of v and add v and (maybe) the sibling of v. Compared to the layer at time $\hat{\tau} \circ \pi_1(\pi_Y \circ p_0^{-1}(v))$, the red degree can only increase because some $Z^*_{R_i}$ from a vertex u are adjacent to v and its sibling. It follows that the maximum $Z^*_{R_i}$ is at most doubled. ◀

6 Ordered twin models

In this section, a transduction pairing of binary tree-orders and preordered rooted binary trees will allow use to consider ordered twin-models (having a sparse reduct) instead of full twin-models.

▶ **Lemma 6.1.** *Let $\mathscr{O}$ be the class of binary tree orders and let $\mathscr{Y}^<$ be the class of rooted binary trees ordered by some preorder. Then there exists simple transductions* L *and* O *such that* (L,O) *is a transduction pairing of $\mathscr{O}$ and $\mathscr{Y}^<$.*

Proof. We define two simple transductions. The first transduction maps binary tree-orders $\prec$ into the preorder defined by some traversal of Y.

L is defined as follows: we consider a mark M on the vertices and define

$$\rho_E(x,y) := \big((x \prec y) \wedge\ \forall v\ \neg(x \prec v \wedge v \prec y)\big) \vee \big((y \prec x) \wedge\ \forall v\ \neg(y \prec v \wedge v \prec x)\big)$$

$$\rho_<(x,y) := (x \prec y) \vee\ \neg(y \preceq x) \wedge \exists u\, \exists v\, \exists w\, \Big(\forall z \big(w \prec z \rightarrow \neg(z \prec u \wedge z \prec v)\big) \wedge (u \preceq x \wedge v \preceq y \wedge w \prec u \wedge w \prec v \wedge \rho_E(u,w) \wedge \rho_E(v,w) \wedge M(u)\Big).$$

Consider a binary tree order $(V(Y),\prec) \in \mathscr{O}$, and let Y be the rooted binary tree defined by $\prec$. Recall that the *preorder* of Y defined by some plane embedding of Y (that is to an ordering, for each node v, of the children of v) is a linear order on $V(Y)$ such that for every internal node v of Y, one finds in the ordering the vertex v, then the first children of v and its descendants, then the second children of v and its descendants. Let $<$ be the preorder defined by some plane embedding of Y. Let us mark by M all the nodes of Y that are the first children of their parent. The formula ρ_E defines the cover graph of $\prec$, thus E is the adjacency relation of Y. Let x,y be nodes of Y. If $x = y$, then $\rho_<(x,y)$ does not hold. If x and y are comparable in $\prec$, then $\rho_<(x,y)$ is equivalent to $x \prec y$. Otherwise, let w be the infimum of x and y, and let u and v be the children of z such that $u \prec x$ and $v \prec y$. Then $\rho_<(x,y)$ holds if u is the first children of w, that is if w is marked. Altogether, we have $\rho_<(x,y)$ if and only if $x < y$. Hence $(V(Y),\prec) \in \mathsf{L}(Y^<)$, where $Y^<$ stands for Y ordered by $<$.

The transduction O is defined as follows:

$$\rho_\prec(x,y) := (x < y) \wedge \forall z \forall w\ (x < z \wedge z \leq y \wedge E(z,w)) \rightarrow (x \leq w).$$

Let $Y^<$ be a rooted binary tree Y with preorder $<$ and let $\prec$ be the corresponding tree-order. If $x \geq y$, then $\rho_\prec(x,y)$ does not hold so we assume $x < y$. Assume x is an ancestor of y in Y, then all the vertices z between x and y in the preorder are descendants of x thus any neighbour of these are either descendants of x or x itself thus $\rho_\prec(x,y)$ holds. Otherwise, let w be the infimum of x and y in Y and let z be the children of z that is an ancestor of y. Then z is between x and y is the preorder, z is adjacent to w, but w appears before x in the preorder. Thus $\rho_\prec(x,y)$ does not hold. It follows that $\rho_\prec(x,y)$ is equivalent to $x \prec y$ thus $Y^< \in \mathsf{O}(V(Y),\prec)$. Thus (L,O) is a transduction pairing of $\mathscr{O}$ and $\mathscr{Y}^<$. ◀

▶ **Lemma 6.2.** *Let $\mathscr{C}$ be a class of Σ-structures with bounded twin-width, and let $\mathscr{T}$ be a class of optimal twin-models of the Σ-structures in $\mathscr{C}$. Then there exist a simple transduction* Unfold *such that* $(\mathsf{Gaifman},\mathsf{Unfold})$ *is a transduction pairing of $\mathscr{T}$ and* $\mathsf{Gaifman}(\mathscr{T})$*.*

Proof. Let $\mathscr{F}$ be the class of full twin-models of the Σ-structures in $\mathscr{C}$ derived from the twin-models in $\mathscr{C}$. According to Lemma 5.2 the class $\mathscr{F}$ has bounded twin-width. Let $\mathscr{T}^<$ be the class of ordered twin-models obtained from the twin-models in $\mathscr{T}$ by adding a linear order given by the preorder of the rooted tree of the tree model. Then $\mathscr{T}^<$ is an L-transduction of $\mathscr{F}$ hence has bounded twin-width. Thus the class $\mathscr{T}$, being a reduct of $\mathscr{T}^<$, has bounded twin-width. It follows from Lemma 4.13 that the class $\mathsf{Gaifman}(\mathscr{T})$ is degenerate, hence, according to Theorem 3.2, it has bounded expansion and, according to Theorem 3.3, bounded star chromatic number. It follows from Lemma 3.1 that $(\mathsf{Gaifman},\mathsf{Unfold})$ is a transduction pairing of $\mathscr{T}$ and $\mathsf{Gaifman}(\mathscr{T})$. ◀

7 Permutations and the main result

When we speak about transductions of permutations, we consider the permutations as defined in Section 2.1. Hence the language used to define the transduction can use the binary relations $<_1, <_2$, as well as equality. Note that the mapping $x \mapsto \sigma(x)$ is not first-order definable in this setting.

Let $\sigma \in \mathfrak{S}_n$. The permutation σ can be represented by the balanced ordered matching $\mathbf{M}_\sigma^< = ([2n], E, <)$, where $<$ is the natural order on $[2n]$ and E is a perfect matching in which $i \in [n]$ is matched with $\sigma(i) + n$.

▷ Claim 7.1. Let $c \in \mathbb{N}$ and let $\mathscr{C}^<$ be a class of ordered graphs with star chromatic number at most c. There exist a copying-transduction T_1 with $\mathrm{bf}(\mathsf{T}_1) = c$, a simple transduction T_2, and a class $\mathscr{P}$ of permutations such that $(\mathsf{T}_1, \mathsf{T}_2)$ is a transduction pairing of $\mathscr{C}^<$ and $\mathscr{P}$.

Proof. Let $\mathscr{C}$ be the reduct of $\mathscr{C}^<$ obtained by "forgetting" the linear order. As $\mathscr{C}$ has star chromatic number at most c, each graph $G \in \mathscr{C}$ has an orientation with indegree at most $c-1$. Let Σ_c be the relational signature consisting of $c-1$ binary relations $E_1, \dots, E_{c-1}$ and let $\Sigma_c^<$ be the signature obtained from Σ_c by adding a linear order relation $<$. By dispatching the incoming edges at each vertex, we associate to each $G \in \mathscr{C}$ an asymmetric Σ_c-structure $\mathbf{A}_G$ such that $G = \mathsf{Gaifman}(\mathbf{A}_G)$ and every vertex of $\mathbf{A}_G$ has indegree at most 1 in each relation. We also associate to each $G^< \in \mathscr{C}^<$ a $\Sigma_c^<$-structure $\mathbf{A}_G^<$ by copying the linear order of $G^<$. Let $\mathscr{D} = \{\mathbf{A}_G : G \in \mathscr{C}\}$ and let $\mathscr{D}^< = \{\mathbf{A}_G^< : G \in \mathscr{C}^<\}$. According to Lemma 3.1, there exists a transduction Unfold such that $(\mathsf{Unfold}, \mathsf{Gaifman})$ is a transduction pairing of $\mathscr{C}$ and $\mathscr{D}$. This transduction also defines a transduction pairing $(\mathsf{T}_U, \mathsf{T}_G)$ of $\mathscr{C}^<$ and $\mathscr{D}^<$.

The copying-transduction T_1 is the composition of T_B, a c-blowing, and the transduction defined as follows: Let f_i be the first-order definable function that maps a vertex x to its in-neighbor in relation E_i if it exists, and to x, otherwise. We further define f_c as the identity. We define $<_1$ as the lexicographic order and $<_2$ as follows: $(x, i) <_2 (y, j)$ if $f_i(x) < f_j(y)$ or $f_i(x) = f_j(y)$ and $i < j$, or $f_i(x) = f_j(y), i = j$, and $x < y$.

The definition of T_2 is as follows: we partition V into $V_1, \dots, V_c$ in such a way that the maximum element of $<_1$ is in V_c. The linear order $<$ is the restriction of $<_1$ to V_c. Then we define the mapping $g_i : V_c \to V_c$ as follows: for $x \in V_c$, if there exists no $x' \in V_i$ such that $x' <_1 x$ and no element between x' and x (in $<_1$) belongs to $V_i \cup V_c$ then $g_i(x) = x$. Otherwise, $g_i(x)$ is, with respect to $<_2$, the minimum element of V_c greater than x'.

Let $f_1, \dots, f_{c-1}$ be mappings on V and let $<_1$ and $<_2$ be defined by T_1. We let $V_i = V \times \{i\}$. By definition, the maximum element of $V \times [c]$ with respect to $<_1$ is in V_c. Let $i \in [c-1]$ and $(x, c) \in V_c$. By construction, $x' = (x, i)$, and (in $<_2$) the minimum element in V_c greater than x' is $(f_i(x), c)$. Thus $g_i((x, c)) = (f_i(x), c)$. Hence $\mathsf{T}_2 \circ \mathsf{T}_1$ subsumes the identity. It follows that $(\mathsf{T}_1, \mathsf{T}_2)$ is a transduction pairing of $(\mathscr{C}^<)$ and $\mathscr{P} = \mathsf{T}_1(\mathscr{C}^<)$. ◁

▶ **Theorem 7.2.** *For every class $\mathscr{C}$ of binary structures with twin-width at most t there exists a proper class of permutations $\mathscr{P}$, an integer k, and a transduction T from $\mathscr{P}$ to $\mathscr{C}$, such that for every graph $G \in \mathscr{C}$ there is a permutation $\sigma \in \mathscr{P}$ on $k|G|$ elements with $G \in \mathsf{T}(\sigma)$.*

Proof. Let $\mathscr{C}$ be a class of binary structures with twin-width at most t. Let $\mathscr{T}$ be a class of twin-models obtained by optimal contraction sequences of graphs in $\mathscr{C}$, and let $\mathscr{F}$ be the class of the corresponding full twin-models. According to Lemma 5.2 $\mathscr{F}$ has twin-width at most $2t$, moreover, applying the transduction L on $\prec$ we transform $\mathscr{F}$ into the class $\mathscr{T}^<$, whose reduct is $\mathscr{T}$. Let $\mathscr{D}^<$ be the class obtained from $\mathscr{T}^<$ be taking the Gaifman graphs of the relations distinct from the linear order, and keeping the linear order, and let $\mathscr{D}$ be the reduct of $\mathscr{D}^<$

obtained by forgetting the linear order. Thus $\mathscr{D} = \mathsf{Gaifman}(\mathscr{T})$. As the classes $\mathscr{D}^<$ and its reduct $\mathscr{D}$ are transductions of the class $\mathscr{F}$ they have bounded twin-width. Moreover, the class $\mathscr{D}$ is degenerate hence it has bounded expansion and, in particular, bounded star chromatic number. It follows that we have a transduction pairing of $\mathscr{D}^<$ and a class $\mathscr{P}$ of permutations, which is proper as it has bounded twin-width. From the transduction pairing of $\mathscr{T}^<$ and $\mathscr{D}^<$ and the one of $\mathscr{F}$ and $\mathscr{T}^<$ we deduce that there is a transduction pairing of $\mathscr{F}$ and $\mathscr{P}$. As $\mathscr{C}$ is a transduction of $\mathscr{F}$ we conclude that $\mathscr{C}$ is a transduction of $\mathscr{P}$. ◀

Note that any transduction from the class $\mathscr{P}$ of permutations to the class $\mathscr{C}$ of Σ-structures obviously defines a transduction from the permutation class $\overline{\mathscr{P}}$ obtained by closing $\mathscr{P}$ under sub-permutations to the class $\mathscr{C}$.

▶ **Corollary 7.3.** *Every class of graphs with bounded twin-width contains at most c^n non-isomorphic graphs on n vertices (for some constant c depending on the class).*

8 Further remarks

The *growth constant* of a class $\mathscr{C}^{\rm lab}$ of labeled graphs is defined as $\gamma_{\mathscr{C}}^{\rm lab} = \limsup(|\mathscr{C}_n^{\rm lab}|/n!)^{1/n}$, where $\mathscr{C}_n$ denotes the set of all graphs in $\mathscr{C}$ which have n vertices. By analogy, the *unlabeled growth constant* of a class $\mathscr{C}$ of (unlabeled) graphs is defined as $\gamma_{\mathscr{C}} = \limsup|\mathscr{C}_n|^{1/n}$. Bounding the star chromatic number of a twin-model of a graph with twin-width d as a function of d would allow to give some upper bound on the constant $\gamma_{\mathscr{C}}$ for a class $\mathscr{C}$ with bounded twin-width.

It has been conjectured in [4] that every small hereditary class of graphs has bounded twin-width. According to Corollary 7.3 this is equivalent to the following conjecture.

▶ **Conjecture 8.1.** *For a hereditary class of graphs $\mathscr{C}$ the following are equivalent:*

1. *$\mathscr{C}$ has bounded twin-width;*
2. *$\mathscr{C}$ contains at most c^n non-isomorphic graphs with n vertices, for some constant c;*
3. *$\mathscr{C}$ is small, that is: $\mathscr{C}$ contains at most $c^n\,n!$ labeled graphs with n vertices, for some constant c.*

It was announced by Simon and Toruńczyk [19], and independently proved by Bonnet *et al.* [5] that a class of graphs $\mathscr{C}$ has bounded twin-width if and only if it is the reduct of a monadically dependent class of ordered graphs. This implies the following duality type statement for every class $\mathscr{C}^<$ of ordered graphs:

$$\exists\sigma\ (\ \mathrm{Av}(\sigma) \longrightarrow\!\!\!\!\!\rightarrow \mathscr{C}^<\ ) \qquad \iff \qquad \mathscr{C}^< \not\longrightarrow\!\!\!\!\!\rightarrow \mathscr{G}\ ,$$

where $\mathrm{Av}(\sigma)$ denotes the class of all permutations avoiding the pattern σ and $\mathscr{G}$ denote the class of all graphs.

The connection between classes of ordered graphs and permutation classes might well be even deeper than what is proved in this paper.

▶ **Conjecture 8.2.** *Every hereditary class $\mathscr{C}^<$ of ordered graphs is transduction equivalent to a permutation class.*

This conjecture is known to hold in the following cases:

- if the class $\mathscr{C}^<$ is not monadically dependent, as it is then transduction equivalent to the class of all permutations [5];
- if the reduct $\mathscr{C}$ of $\mathscr{C}^<$ is biclique-free, as either $\mathscr{C}^<$ is not monadically dependent (previous item), or it has bounded twin-width [5] thus, according to Theorem 3.2, Theorem 3.3 and Claim 7.1 the class $\mathscr{C}^<$ is transduction equivalent to a permutation class;
- if the reduct $\mathscr{C}$ of $\mathscr{C}^<$ is a transduction of a class with bounded expansion as $\mathscr{C}$ is then transduction equivalent to a bounded expansion class $\mathscr{D}$ [10] and this transduction equivalence can be extended to a transduction equivalence of $\mathscr{C}^<$ and an expansion $\mathscr{D}^<$ of $\mathscr{D}$, which falls in the second item.

References

1 Michael Albert, Mathilde Bouvel, and Valentin Féray. Two first-order logics of permutations. *Journal of Combinatorial Theory, Series A*, 171:105158, 2020.

2 Olivier Bernardi, Marc Noy, and Dominic Welsh. Growth constants of minor-closed classes of graphs. *Journal of Combinatorial Theory, Series B*, 100(5):468–484, 2010. doi:10.1016/j.jctb.2010.03.001.

3 Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width III: Max Independent Set and Coloring. *CoRR*, abs/2007.14161, 2020. URL: https://arxiv.org/abs/2007.14161, arXiv:2007.14161.

4 Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width II: small classes. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1977–1996, 2021. doi:10.1137/1.9781611976465.118.

5 Édouard Bonnet, Ugo Giocanti, Patrice Ossona de Mendez, and Stéphan Thomassé. Twin-width IV: low complexity matrices. arXiv:2102.03117 [math.CO], 2021.

6 Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width I: tractable FO model checking. In *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 Peter J. Cameron. Homogeneous permutations. *The Electronic Journal of Combinatorics*, page R2, 2002.

8 Bruno Courcelle. The monadic second-order logic of graphs I: recognizable sets of finite graphs. *Inform. and Comput.*, 85:12–75, 1990.

9 Bruno Courcelle. The monadic second order logic of graphs VI: On several representations of graphs by relational structures. *Discrete Applied Mathematics*, 54(2-3):117–149, 1994.

10 Jakub Gajarský, Stephan Kreutzer, Jaroslav Nešetřil, Patrice Ossona de Mendez, Michał Pilipczuk, Sebastian Siebertz, and Szymon Toruńczyk. First-order interpretations of bounded expansion classes. *ACM Transactions on Computational Logic*, 21(4):Article 29, 2020. doi:10.1145/3382093.

11 Sylvain Guillemot and Dániel Marx. Finding small patterns in permutations in linear time. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 82–101, 2014. doi:10.1137/1.9781611973402.7.

12 Wilfrid Hodges et al. *A shorter model theory.* Cambridge university press, 1997.

13 Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley-Wilf conjecture. *J. Comb. Theory, Ser. A*, 107(1):153–160, 2004. doi:10.1016/j.jcta.2004.04.002.

14 Jaroslav Nešetřil and Patrice Ossona de Mendez. Grad and classes with bounded expansion I. Decompositions. *European Journal of Combinatorics*, 29(3):760–776, 2008. doi:10.1016/j.ejc.2006.07.013.

15 Jaroslav Nešetřil, Patrice Ossona de Mendez, Roman Rabinovich, and Sebastian Siebertz. Linear rankwidth meets stability. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms*, pages 1180–1199, 2020. URL: `https://epubs.siam.org/doi/abs/10.1137/1.9781611975994.72`, `doi:10.1137/1.9781611975994.72`.

16 László Rédei. Ein kombinatorischer Satz. *Acta Litt. Szeged*, 7(39-43):97, 1934.

17 Neil Robertson and Paul D. Seymour. Graph minors I–XXIII. 1983–2010.

18 Benjamin Rossman. Homomorphism preservation theorems. *J. ACM*, 55(3):1–53, 2008. `doi:10.1145/1379759.1379763`.

19 Pierre Simon and Szymon Toruńczyk. Ordered graphs of bounded twin-width. personal communication, 2021.