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

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

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) and from the German Research Foundation (DFG) with grant agreement No 444419611.

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 treedepth, treewidth, cliquewidth, shrubdepth 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 [21], Courcelle's theorem on MSO definable properties [10, 11], Robertson and Seymour's graph minor theory [20], and low treedepth decompositions of classes with bounded expansion [16].

42nd Conference on Very Important Topics (CVIT 2016).
Editors: John Q. Open and Joan R. Access; Article No. 23; pp. 23:1–23:22
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

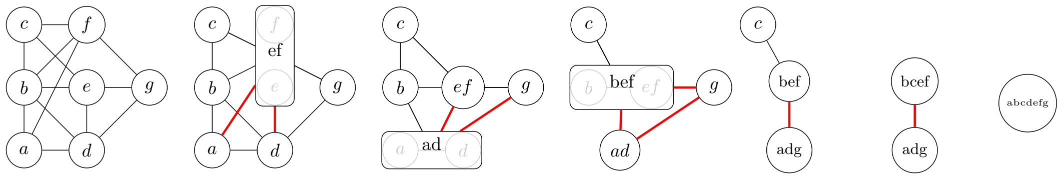


Figure 1 A 2-sequence witnessing that the initial graph has twin-width at most 2.

In this paper we consider the twin-width graph parameter, defined by Bonnet, Kim, Thomassé and Watrigant [7] as a generalization of a width invariant for classes of permutations defined by Guillemot and Marx [13]. This parameter was intensively studied recently in the context of many structural and algorithmic questions such as FPT model checking [7], graph enumeration [4], graph coloring [5], and structural properties of matrices and ordered graphs [6]. The twin-width of graphs is originally defined using a sequence of near-twin vertex contractions or identifications. Roughly speaking, twin-width measures the accumulated error (recorded via the so-called *red edges*) made by the identifications. To help the reader start forming intuitions, we give a concise informal definition of the twin-width of a graph; a formal generalization for binary structures is presented in Section 2.4. A *trigraph* is a graph with some edges colored red (while the rest of them are black). A *contraction* (or identification) consists of merging two (non-necessarily adjacent) vertices, say, u, v into a vertex w, and keeping every edge wz black if and only if uz and vz were black edges. The other edges incident to w turn (or already are) red, and the rest of the trigraph does not change. A *contraction sequence* of an n-vertex graph G is a sequence of trigraphs $G = G_n, \ldots, G_1$ such that G_i is obtained from G_{i+1} by performing one contraction (observe that G_1 is the 1-vertex graph). A *d-sequence* is a contraction sequence where all the trigraphs have red degree at most d. The *twin-width* of G is then the minimum integer d such that G admits a d-sequence. See Figure 1 for an example of a graph admitting a 2-sequence. In the current paper, the definition of twin-width for binary relational structures perfectly matches the one in [7] on undirected graphs, but will slightly differ in the generality of binary structures. Though, as we will observe, both definitions are linearly tied.

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 a ranking, 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. Note that an informal precursor of ranked twin-models appear in [5] in the form of the so-called *ordered union tree* and the realization that the edge set of graphs of twin-width at most d can be partitioned into $O_d(n)$ bicliques where both sides of each biclique is a discrete interval along a unique fixed vertex ordering. The main novelty in the (ranked) twin-models lies in the axiomatization of *legal* sets of transversal edges, which is indispensable to their logical treatment.

This paper is a combination of model-theoretic tools (relational structures, interpretations, transductions), structural graph theory and theory of permutations. The following is the main result of this paper:

▶ **Theorem.** *A class of binary relational structures has bounded twin-width if and only if it is a transduction of a proper permutation class.*

We recall that a proper permutation class is a set of permutations closed under sub-permutations that excludes at least one permutation.

(This is stated in more technical terms in Section 7 as Theorem 7.2.) The fact that any class of graphs with bounded twin-width is just a transduction of a very simple class (such as proper permutation classes) 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 [6]. It can also be thought of as scaling up the fact that classes of bounded rank-width coincide with transductions of tree orders, and classes of bounded linear rank-width, with transductions of linear orders [9]. On the other hand twin-models are interesting objects per se and in a way present one of the most permissive forms of width parameters related to trees. Note that for other classes of sparse structures we do not have such a concrete model.

As the transduction involved in the main result is k-bounded (meaning that each graph on n vertices can be encoded in a permutation on at most kn elements), it is then a consequence of [15] 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 twin-width" machinery (but only the preservation of bounded twin-width by transduction proved in [7]). 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]).

The proof of our main result is surprisingly complex and proceeds in several steps which perhaps add new aspects to the rich spectrum of structures related to twin-width. The basic steps can be outlined as follows (the relevant terminology will be formally introduced in the appropriate sections).

We start with a class $\mathscr{C}_0$ 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 $\mathscr{C}_0$ 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 the class of full twin-models $\mathscr{F}$ 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 thus it has a bounded twin-width) we prove a transduction pairing of $\mathscr{T}$ and $\mathscr{G}$, from which we derive a transduction pairing of $\mathscr{T}^<$ and the class $\mathscr{G}^<$ of ordered Gaifman graphs of the ordered twin-models. As a composition of a transduction pairing of $\mathscr{G}^<$ with a class $\mathscr{E}^<$ of ordered binary structures, in which each binary relation induces a pseudoforest and a transduction pairing of $\mathscr{E}^<$ with a class $\mathscr{P}$ of permutations we define a transduction pairing of $\mathscr{G}^<$ and $\mathscr{P}$. As $\mathscr{G}^<$ 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 $\mathscr{C}_0$ is a transduction of the hereditary closure $\overline{\mathscr{P}}$ of $\mathscr{P}$, which is a proper permutation class.

This proof may be schematically outlined by Fig. 1. Here all the notations are consistent with the notation used later in our proof.

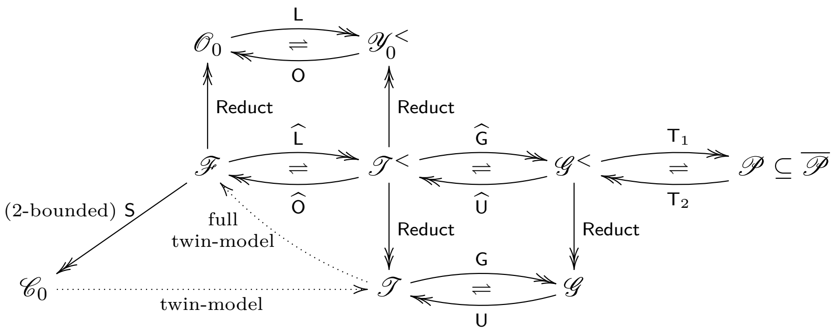


Figure 2 Relations between the classes of structures involved in the proof of the main result. The class $\mathscr{T}$ is defined in Definition 4.5, the class $\mathscr{F}$ and the interpretation S in Definition 5.1, the transduction pairing (L, O) in Lemma 6.1, the classes $\mathscr{O}_0$, $\mathscr{Y}_0^<$, and $\mathscr{T}^<$ in Definition 6.2 (and the transduction pairing $(\widehat{\mathsf{L}}, \widehat{\mathsf{O}})$ as a remark just after), the class $\mathscr{G}$ in Definition 6.3, the transduction pairing (G, U) in Lemma 6.4, the class $\mathscr{G}^<$ in Definition 6.5 (and the transduction pairing $(\widehat{\mathsf{G}}, \widehat{\mathsf{U}})$ as a remark just after).

2 Preliminaries

2.1 Relational structures

We assume basic knowledge of first-order logic and refer to [14] 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

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

(In this paper, by formula, we mean a first-order formula in the language of Σ-structures, where Σ is usually understood from the context.) 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 irreflexive 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 $V(Y) = I(Y) \cup L(Y)$ the set of vertices 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* (also called *total 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. An element x is *covered* by an element y if $x \prec y$ and there is no element z with $x \prec z \prec y$. A *binary tree order* is a tree order such that every non-maximal element is covered by exactly two elements.

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. [8, 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 (see [3]).

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 $\rho_{R'}(x_1, \ldots, x_k)$ for each k-ary relation symbol $R' \in \Sigma'$. Let I be an interpretation of Σ'-structures in Σ-structure, where $\Sigma' = \{R_1', \ldots, R_s'\}$. For each Σ-structure $\mathbf{A}$ we denote by $\mathsf{I}(\mathbf{A}) = (\rho_0(\mathbf{A}), \rho_{R_1'}(\mathbf{A}), \ldots, \rho_{R_s'}(\mathbf{A}))$ 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 $\rho_R(x_1, \ldots, x_k) = R(x_1, \ldots, x_k)$ 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 $k \in \mathbb{N}$. The *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 $\mathbf{A} \models R(p(x_1), \dots, p(x_k))$. 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 $\mathscr{C} \longrightarrow\!\!\!\!\!\!\!\!\!\twoheadrightarrow \mathscr{D}$, if there exists a transduction T such that $\mathscr{D}$ is a T-transduction of $\mathscr{C}$. The negation of $\mathscr{C} \longrightarrow\!\!\!\!\!\!\!\!\!\twoheadrightarrow \mathscr{D}$ is denoted by $\mathscr{C} \not\twoheadrightarrow \mathscr{D}$. Note that we require only the inclusion of $\mathscr{D}$ in $\mathsf{T}(\mathscr{C})$, and not the equality. 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 (D, C) is a transduction pairing, then $\mathscr{D}$ is $\mathrm{bf}(\mathsf{C})$-bounded D-transduction of $\mathscr{C}$ and $\mathscr{C}$ is $\mathrm{bf}(\mathsf{D})$-bounded C-transduction of $\mathscr{D}$. The following easy lemma will be useful.

▶ **Lemma 2.1.**

Assume $\mathscr{D}$ is a D-transduction of $\mathscr{C}$ and for every $\mathbf{A} \in \mathscr{C}$ and every $\mathbf{B} \in \mathsf{D}(\mathbf{A}) \cap \mathscr{D}$ we have $\mathbf{A} \in \mathsf{C}(\mathbf{B})$. Then (D, C) is a transduction pairing of $\mathscr{C}$ and $\mathscr{D}$.

Proof. Let $\mathbf{B} \in \mathscr{D}$. As $\mathscr{D}$ is a D transduction of $\mathscr{C}$ there exists $\mathbf{A} \in \mathscr{C}$ with $\mathbf{B} = \mathsf{D}(\mathbf{A})$. Then $\mathbf{A} \in \mathsf{C}(\mathbf{B}) \cap \mathscr{C}$. ◀

Note that a transduction T from Σ_1-structures to Σ_2-structure naturally defines a transduction $\widehat{\mathsf{T}}$ from Σ_1^+-structures to Σ_2^+-structures if $\Sigma_2^+ \setminus \Sigma_2 = \Sigma_1^+ \setminus \Sigma_1$ by leaving the relations in $\Sigma_1^+ \setminus \Sigma_1$ unchanged. The transduction $\widehat{\mathsf{T}}$ is called the *natural generalization* of T to Σ_1^+-structures.

2.4 Twin-width

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 [7]. 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 [7].

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 $\mathbf{A} \models R(w, u)$, and $\mathbf{A}' \models R(z, w)$ if $\mathbf{A} \models R(u, w)$ (Note that this does not change if we use v instead of 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 $\mathbf{A}_n, \dots, \mathbf{A}_1$ 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 $1 \le i < n$, $\mathbf{A}_i$ is obtained from $\mathbf{A}_{i+1}$ by performing a single contraction; for every $1 \le i < n$ 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 v in the undirected graph $(A, R^*(\mathbf{A}))$. When d is not specified, we shall speak of a *contraction sequence*. The minimum d such that there exists a d-sequence for a Σ-structure $\mathbf{A}$ is the *twin-width* $\operatorname{tww}(\mathbf{A})$ of $\mathbf{A}$. This definition for binary relational structures differs from the one given in [7] (where red edges are not counted with multiplicity), but will be more convenient in our setting. However, the definitions differ by at most a constant factor (linear in $|\Sigma|$), thus the derived notions of class with bounded twin-width coincide. A crucial property of twin-width is the following result.

▶ **Theorem 2.2** ([7, Theorem 7]). *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 Classes with bounded star chromatic number

One of the key ingredient of the proof will rely on a pairing between a class $\mathscr{C}$ and the class of the Gaifman graphs of the structures in $\mathscr{C}$. Though such a pairing does not exist for general classes of structures, we prove in this section that this is the case if the Gaifman graphs have bounded star chromatic number.

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* $\chi_{\rm st}(G)$ of G is the minimum number of colors in a star coloring of G. Note that a star coloring of a graph with c colors 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 $\mathsf{Unfold}_{\Sigma,c}$ from graphs to Σ-structures such that if the Gaifman graphs of the structures in $\mathscr{C}$ have star chromatic number at most c, then $(\mathsf{Gaifman}_\Sigma, \mathsf{Unfold}_{\Sigma,c})$ is a transduction pairing of $(\mathscr{C}, \mathsf{Gaifman}_\Sigma(\mathscr{C}))$.*

Proof. Let $c = \sup\{\chi_{\mathrm{st}}(G) : G \in \mathsf{Gaifman}_\Sigma(\mathscr{C})\} < \infty$. Let $\mathbf{A} \in \mathscr{C}$, let $G = \mathsf{Gaifman}_\Sigma(\mathbf{A})$, 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 neighbor of v with color $\gamma(u)$. Let $R \in \Sigma$ be a relation of arity k. For each $(u_1, \dots, u_k) \in R(\mathbf{A})$, $u_1, \dots, u_k$ induce a tournament in $\vec{G}$. Every tournament has at least one directed Hamiltonian path [19]. We fix one such Hamiltonian path and let $p(u_1, \dots, u_k)$ be the index of the last vertex in the path. Let $a = p(u_1, \dots, u_k)$, let $(c_1, \dots, c_k) = (\gamma(u_1), \dots, \gamma(u_k))$. Then there exists in G exactly one clique of size k containing u_a with vertices colored $c_1, \dots, c_k$. Indeed, as $\mathbf{A} \models R(u_1, \dots, u_k)$, the vertices of $u_1, \dots, u_k$ induce a clique in G. As for $i \neq a$ the vertex u_i is an in-neighbors of u_a in $\vec{G}$, it is, in G, the only neighbor of u_a with color c_i. Thus no other clique of size k with vertices colored $c_1, \dots, c_k$ can contain u_a. For each relation $R \in \Sigma$ with arity k and each $(u_1, \dots, u_k) \in R(\mathbf{A})$ we put at $v = u_{p(u_1,\dots,u_k)}$ a mark $M^R_{\gamma(u_1),\dots,\gamma(u_k)}$. Then, in the graph G, the vertex v belongs to exactly one clique of size k with vertices colored $\gamma(u_1), \dots, \gamma(u_k)$, what allows to recover the tuple $(u_1, \dots, u_k)$, as all the colors are distinct. We further put at each vertex v a mark $C_{\gamma(v)}$. Then the structure $\mathbf{A}$ is reconstructed by the transduction $\mathsf{Unfold}_{\Sigma,c}$ defined by the formulas

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

◀

Note that if the condition of Lemma 3.1 is tight in the following sense: if a class $\mathscr{C}$ of undirected graphs contains graphs with arbitrarily large star chromatic number and girth then the class $\vec{\mathscr{C}}$ of all orientations of the graphs in $\mathscr{C}$ is not a transduction of $\mathscr{C}$ [18].

Lemma 3.1 will be particularly significant in conjunction with the following results. Recall that a graph G is *d-degenerate* if every induced subgraph of G contains a vertex of degree at most d, and that a class $\mathscr{C}$ is *degenerate* if all the graphs in $\mathscr{C}$ are d-degenerate, for some d. A class $\mathscr{C}$ has *bounded expansion* if, for every integer k, the class of all graphs H whose k-subdivision is a subgraph of some graph in $\mathscr{C}$ is degenerate [17].

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

▶ **Theorem 3.3** ([16]). *Every class of graphs 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 [5]. 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:*

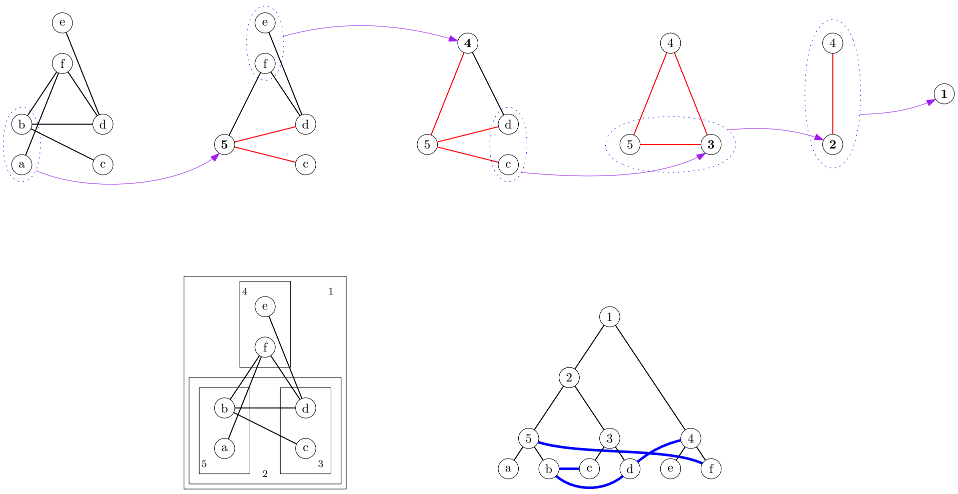


Figure 3 A contraction sequence, a block representation of the contractions, and a twin-model.

- (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, layers, and width)**.** *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*

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

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* $r(Y)$*.*

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

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

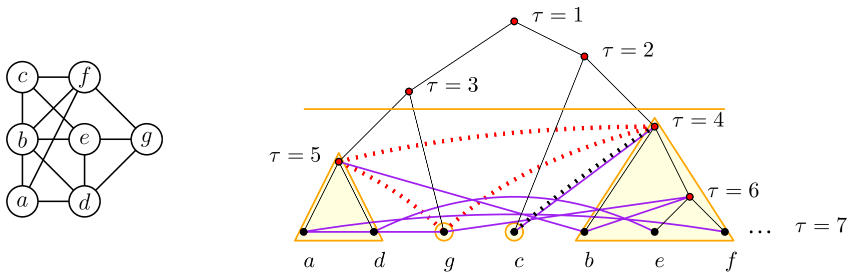


Figure 4 A graph G and a ranked twin-model of G. The boundary $\partial_4 Y$ is the set $\{5, g, c, 4\}$ (internal vertices labeled by τ), which can be represented as the set of the yellow zones. The relations of $\mathbf{L}_4$ are depicted as dotted heavy lines (black for R, red for R^*). The width of this twin-model is 2.

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 5 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.

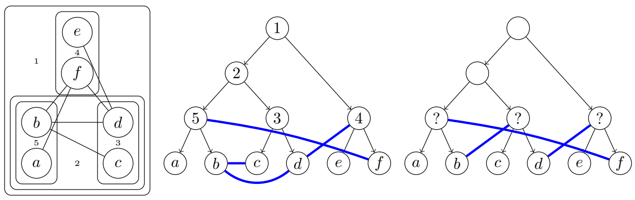


Figure 5 Left: A 6-vertex graph and a contraction sequence, where the tiny digit in each box indicates the index of contracted vertices when they appear. Center: A twin-model of the graph, where the edges of Z_E are in bold blue, and a ranking (for the internal nodes) of this twin-model that actually 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 $5bd4f$ has $(b, d), (d, 4) \in Z_E$.

4.2 From a contraction sequence to a twin-model

In this section we prove that every d-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 4.3** (proved in the appendix). *Every d-sequence $\mathbf{A}_n, \dots, \mathbf{A}_1$ defines a ranked twin-model with width at most d.*

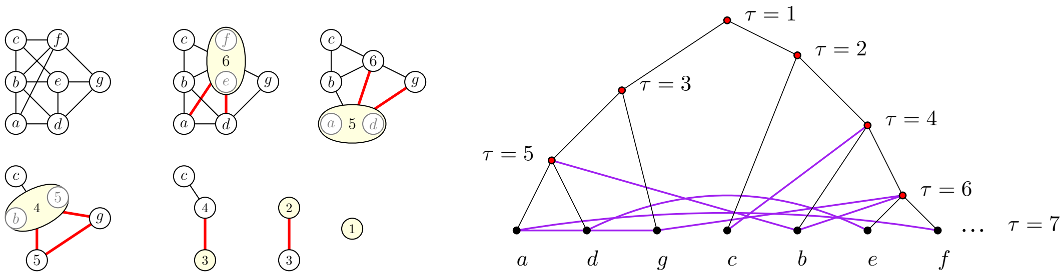


Figure 6 A contraction sequence and the derived ranked tree model.

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 4.4** (proved in the appendix). *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 $\mathbf{A}$.*

Lemma 4.4 allows to introduce the following terminology: the *width* of a twin-model $(Y, Z_{R_1}, \dots, Z_{R_k})$ is the minimum width of a ranking of $(Y, Z_{R_1}, \dots, Z_{R_k})$. A twin-model of a Σ-structure $\mathbf{A}$ is *optimal* is it has the minimum possible width as a twin-model of $\mathbf{A}$, which is the twin-width of $\mathbf{A}$.

▶ **Definition 4.5** (The class $\mathscr{T}$). *The class $\mathscr{T}$ is the class of all optimal twin-models of the Σ-structures in $\mathscr{C}_0$.*

The following easy remark will be useful.

▷ Claim 4.6. 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 pairwise least common 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 4.7.** *The Gaifman graph of a ranked 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 4.4). 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 degree at most $d+1 \leq d+k+1$. 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 $(V(Y), \prec_Y, Z_{R_1}, \dots, Z_{R_k})$.

▶ **Definition 5.1** (Transduction S and the class $\mathscr{F}$). *The transduction* S *is the simple interpretation of* Σ*-structures in full twin-models defined by formulas*

$$\rho_0(x) := \neg(\exists y\ y \succ_Y x);$$
$$\rho_{R_i}(x,y) := \exists u\, \exists v\ (u \preceq_Y x) \wedge (v \preceq_Y y) \wedge Z_{R_i}(u,v).$$

$\mathscr{F}$ is the class of all the full twin-models corresponding to the twin-models in $\mathscr{T}$.

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

▶ **Lemma 5.2.** *The class $\mathscr{C}_0$ is a 2-bounded* S*-transduction of the class* F.

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

▶ **Lemma 5.3** (proved in the appendix). *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 twin-width of $\mathbf{T}$ is at most twice the width of the ranked twin-model $(V(Y), \prec_Y, Z_{R_1}, \dots, Z_{R_k}, \tau)$.*

6 Ordered twin-models

Recall that a *preordering* of the vertices of a rooted tree is the discovery order of the vertices of a (depth-first search) traversal of the tree starting at its root.

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

▶ **Definition 6.2** (The classes $\mathscr{O}_0, \mathscr{Y}_0^<$, and $\mathscr{T}^<$). *The class $\mathscr{O}_0$ is the reduct of the class $\mathscr{F}$, obtained by keeping only the tree order relation; the class $\mathscr{Y}_0$ is the class of all rooted binary trees corresponding to the tree orders in $\mathscr{Y}_0$ ordered by some preordering, so that* (L, O) *is a transduction pairing of $\mathscr{O}_0$ and $\mathscr{Y}_0^<$.*

The class $\mathscr{T}^<$ is the class of ordered twin-models obtained from the twin-models in $\mathscr{T}$ by adding a linear order defined by some preordering of the rooted tree of the tree model.

Note that the natural generalization $(\widehat{\mathsf{L}}, \widehat{\mathsf{O}})$ of (L, O) is a transduction pairing of $\mathscr{F}$ and $\mathscr{T}^<$.

▶ **Definition 6.3** (The class $\mathscr{G}$)**.** *The class $\mathscr{G}$ is the class of the Gaifman graphs of the structures in $\mathscr{T}$.*

▶ **Lemma 6.4.** *Let $\mathsf{G} = \mathsf{Gaifman}_\Sigma$. There exists a simple transduction U such that (G, U) is a transduction pairing of $\mathscr{T}$ and $\mathscr{G}$.*

Proof. According to Lemma 5.3 the class $\mathscr{F}$ has bounded twin-width. According to Theorem 2.2, as $\mathscr{T}^<$ is an L-transduction of $\mathscr{F}$ it has bounded twin-width. Thus the class $\mathscr{T}$, being a reduct of $\mathscr{T}^<$, has bounded twin-width. It follows from Lemma 4.7 that the class $\mathsf{G}(\mathscr{T})$ is degenerate, hence, according to Theorem 3.2, it has bounded expansion and, according to Theorem 3.3, bounded star chromatic number (at most c). It follows from Lemma 3.1 that, defining $\mathsf{U} = \mathsf{Unfold}_{\Sigma,c}$, (G, U) is a transduction pairing of $\mathscr{T}$ and $\mathsf{G}(\mathscr{T}) = \mathscr{G}$. ◀

▶ **Definition 6.5** (The class $\mathscr{G}^<$)**.** *The class $\mathscr{G}^<$ is the class of ordered graphs obtained from the structures in $\mathscr{T}^<$ by applying the natural generalization $\widehat{\mathsf{G}}$ of the interpretation $\mathsf{Gaifman}_\Sigma$.*

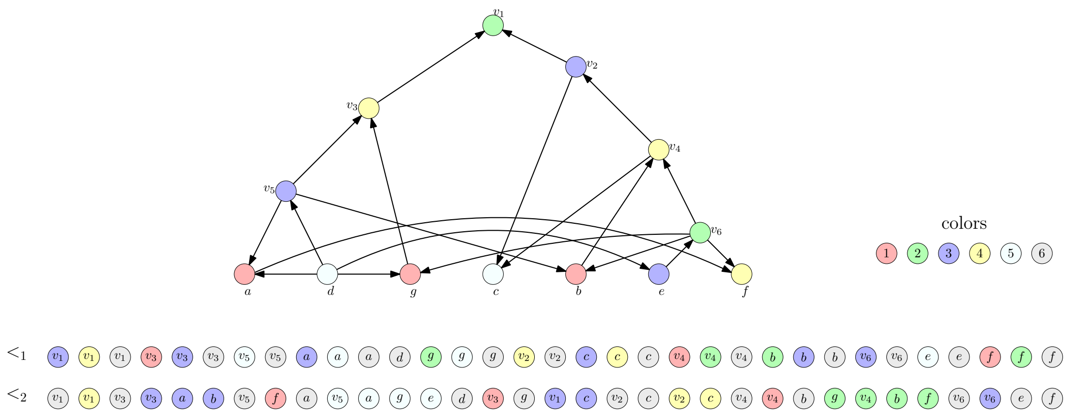
Note that, denoting by $\widehat{\mathsf{U}}$ the natural generalization of the transduction U, it follows from Lemma 6.4 that $(\widehat{\mathsf{G}}, \widehat{\mathsf{U}})$ is a transduction pairing of $\mathscr{T}^<$ and $\mathscr{G}^<$.

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.

▶ **Lemma 7.1** (proved in the appendix)**.** *Let $c \in \mathbb{N}$ and let $\mathscr{G}^<$ be a class of ordered graphs with star chromatic number at most c. There exist a copying-transduction T_1 with $\mathrm{bf}(\mathsf{T}_1) = c + 1$, 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 $\mathscr{G}^<$ and $\mathscr{P}$.*

We give here an informal description of the transductions T_1 and T_2 and refer to Figure 7 for an example: The transduction T_1 is used to compute a permutation σ from an ordered graph $G^<$ and works as follows: we first compute a star coloring γ of G with c colors and orient edges so that bicolored stars are oriented from their centers. Then we blow each vertex into $(u, 1), \ldots, (u, c + 1)$. From this we keep only the vertices of the form $(v, c + 1)$ or of the form (v, i) if v has an in-neighbor colored i. The linear order $<_1$ orders pairs (u, i) by first coordinate first (using $<$) then by increasing i. The linear order < 2 is a succession of intervals ending with a vertex of the form $(v, c + 1)$ (these intervals being ordered according to the order on v); the interval ending with $(v, c + 1)$ contains the pairs $(u, \gamma(v))$, for all the in-neighbors u of v, ordered by first coordinate. The transduction T_2 is used to compute an ordered graph $G^<$ from a permutation σ. It works as follows. First, we mark some elements. These elements will correspond to the vertices of $G^<$, and the linear order $<$ will be the restriction of $<_1$ to these elements. Each vertex v defines a maximal interval $A(v)$ in $<_1$ ending with v and containing no other marked element, and a maximal interval $B(v)$ in $<_1$ ending with v and containing no other marked element. In $G^<$, a vertex u is adjacent to a vertex $v \neq u$ if $A(u)$ intersects $B(v)$ or $A(v)$ intersects $B(u)$.



Thus the permutation obtained for this example is

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 & 25 & 26 & 27 & 28 & 29 & 30 & 31 & 32 & 33 \\ 3 & 2 & 6 & 5 & 9 & 25 & 8 & 31 & 11 & 7 & 10 & 14 & 29 & 12 & 4 & 15 & 1 & 18 & 17 & 20 & 16 & 19 & 23 & 21 & 26 & 13 & 22 & 24 & 32 & 28 & 27 & 30 & 33 \end{pmatrix}$$

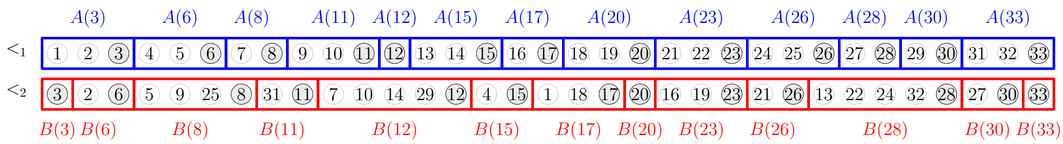


Figure 7 In the top, the transduction T_1 (we assume $c = 5$). In the bottom, the transduction T_2. In gray, the marked elements of σ, which are the vertices of $G^<$. In $G^<$, an edge links 15 (g in the top) and 28 (v_6 in the top) as $A(15)$ intersects $B(28)$.

▶ **Theorem 7.2.** *For every class $\mathscr{C}_0$ of binary structures with twin-width at most t there exists a proper permutation class $\overline{\mathscr{P}}$, an integer k, and a transduction T, such that $\mathscr{C}_0$ is a k-bounded T-transduction of $\overline{\mathscr{P}}$. Precisely, for every graph $G \in \mathscr{C}_0$ there is a permutation $\sigma \in \overline{\mathscr{P}}$ on at most $k|G|$ elements with $G \in \mathsf{T}(\sigma)$.*

Proof. Let $\mathscr{C}_0$ 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 5.3 $\mathscr{F}$ 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 $\mathscr{T}$ (see Lemma 6.1). Let $\mathscr{G}^<$ 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 $\mathscr{G}$ be the reduct of $\mathscr{G}^<$ obtained by forgetting the linear order. Thus $\mathscr{G} = \mathsf{Gaifman}(\mathscr{T})$. As the classes $\mathscr{G}^<$ and its reduct $\mathscr{G}$ are transductions of the class $\mathscr{F}$ they have bounded twin-width, by Theorem 2.2. Moreover, the class $\mathscr{G}$ is degenerate hence it has bounded expansion and, in particular, bounded star chromatic number. It follows that we have a transduction pairing of $\mathscr{G}^<$ and a class $\mathscr{P}$ of permutations, which is proper as it has bounded twin-width. From the transduction pairing of $\mathscr{T}^<$ and $\mathscr{G}^<$ and the one of $\mathscr{F}$ and $\mathscr{T}^<$ we deduce that there is a transduction pairing of $\mathscr{F}$ and $\mathscr{P}$. As $\mathscr{C}_0$ is a transduction of $\mathscr{F}$ we conclude that $\mathscr{C}_0$ is a transduction of $\mathscr{P}$.

Note that the class $\mathscr{C}_0$ is obviously also a T-transduction of the permutation class $\overline{\mathscr{P}}$ obtained by closing $\mathscr{P}$ under sub-permutations. ◀

▶ **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).*

Proof. Let $\mathscr{C}_0$ be a class with bounded twin-width. As twin-width is monotone with respect to induced subgraph inclusion, we may assume that $\mathscr{C}_0$ is hereditary. According to Theorem 7.2, there exists proper permutation class $\overline{\mathscr{P}}$, an integer k, and a transduction T, such that for every $G \in \mathscr{C}_0$ there is a permutation $\sigma \in \overline{\mathscr{P}}$ on $k|G|$ elements with $G \in \mathsf{T}(G)$. Let m be the number of unary predicates used by the transduction. According to Marcus-Tardos theorem [15], for every proper permutation $\mathscr{P}$ there exists a constant a such that $\mathscr{P}$ contains at most a^n permutations on n elements. For each permutation on n elements there are 2^{mn} possible choice for the interpretation of the m predicates (as each predicate defines a subset of elements). It follows that $\mathscr{C}_0$ contains at most $\sum_{i=1}^{kn} a^i\,2^{mi} = O((a^k 2^{mk})^n)$ non-isomorphic graphs with at most n vertices. Thus there exists a constant c such that $\mathscr{C}_0$ contains at most c^n non-isomorphic graphs with n vertices. ◀

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 was proved by Bonnet *et al.* [6] 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:

$$\exists \text{ permutation } \sigma \text{ with } \operatorname{Av}(\sigma) \twoheadrightarrow \mathscr{C}^< \qquad \iff \qquad \mathscr{C}^< \not\twoheadrightarrow \mathscr{U}\ ,$$

where $\operatorname{Av}(\sigma)$ denotes the class of all permutations avoiding the pattern σ and $\mathscr{U}$ 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 8.1.** *Every hereditary class $\mathscr{C}^<$ of ordered graphs is transduction equivalent to a permutation class.*

This conjecture is known to hold if the class $\mathscr{C}^<$ is not monadically dependent, as it is then transduction equivalent to the class of all permutations [6]; it also holds 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 [6]. Then, since biclique-free classes of bounded twin-width have bounded expansion [4], and according to Theorem 3.3 and Lemma 7.1 the class $\mathscr{C}^<$ is transduction equivalent to a permutation class; finally, it also holds 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 $\mathscr{D}$ [12] 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 previous case.

Acknowledgments

The authors wish to express their gratitude to the referees of a first version of this paper, for their careful reading and their most valuable suggestions.

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 Miklós Bóna. *Combinatorics of permutations*. CRC Press, 2012.

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, 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. doi:10.4230/LIPIcs.ICALP.2021.35.

6 Édouard Bonnet, Ugo Giocanti, Patrice Ossona de Mendez, Pierre Simon, Stéphan Thomassé, and Szymon Toruńczyk. Twin-width IV: ordered graphs and matrices, 2021. arXiv:2102.03117.

7 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.

8 Peter J. Cameron. Homogeneous permutations. *The Electronic Journal of Combinatorics*, page R2, 2002.

9 Thomas Colcombet. A combinatorial theorem for trees. In Lars Arge, Christian Cachin, Tomasz Jurdzinski, and Andrzej Tarlecki, editors, *Automata, Languages and Programming, 34th International Colloquium, ICALP 2007, Wroclaw, Poland, July 9-13, 2007, Proceedings*, volume 4596 of *Lecture Notes in Computer Science*, pages 901–912. Springer, 2007. doi:10.1007/978-3-540-73420-8_77.

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

11 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.

12 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.

13 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.

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

15 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.

16 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.

17 Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity (Graphs, Structures, and Algorithms)*, volume 28 of *Algorithms and Combinatorics*. Springer, 2012. 465 pages.

18 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`.

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

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

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

A Proofs of the lemmas

▶ **Lemma 4.3.** *Every d-sequence $\mathbf{A}_n, \ldots, \mathbf{A}_1$ defines a ranked twin-model with width at most d.*

Proof.

▷ Claim A.1. 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 A.2. 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 A.3. The ranked twin-model $(Y, Z_{R_1}, \ldots, Z_{R_k}, \tau)$ derived from a d-sequence $\mathbf{A}_n, \ldots, \mathbf{A}_1$ has width at most d.

◀

▶ **Lemma 4.4.** *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 A.4. 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 A.5. Let $\mathbf{L}_1, \ldots, \mathbf{L}_n$ be the layers of a ranked tree model $\mathfrak{T}$ of a Σ-structure $\mathbf{A}$. Then there exists a contraction sequence $\mathbf{A}_n, \ldots, \mathbf{A}_1$ 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 and only 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 is equivalent to the fact 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 A.4 it follows that $(z, w_i) \in \overline{R}(\mathbf{A}_i) \iff \{(z, w_i), (w_i, z)\} \subseteq \overline{R}(\mathbf{L}_i) \iff (z, w_i) \in \overline{R}(\mathbf{L}_i)$, thus $\overline{R}(\mathbf{A}_i) = \overline{R}(\mathbf{L}_i)$.

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 A.6. 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 cycle. Replace each arc of the form $\pi(u)v$ of this directed cycle (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 $\bigcup_{R\in\Sigma} Z_R$. We show that we can also find a directed cycle with this property, contradicting the consistency assumption. Consider a shortest closed walk $W = (e_1, \dots, e_m)$ with the above property and assume this closed walk is not a directed cycle. Without loss of generality we can assume that $(e_1, \dots, e_k)$ forms a cycle γ (starting the closed walk at another point if necessary). By minimality of the closed walk, the cycle γ contains two consecutive edges in $\bigcup_{R\in\Sigma} Z_R$. These edges are the edges e_k and e_1 (as otherwise they would be consecutive in W as well). It follows that e_{k+1} does not belong to $\bigcup_{R\in\Sigma} Z_R$ (as it follows e_k in the W). The closed walk $W' = (e_{k+1}, \dots, e_n)$ does not not have two consecutive edges in $\bigcup_{R\in\Sigma} Z_R$ as all the consecutive pairs are consecutive in W, except the pair (e_n, e_{k+1}) (and we know $e_{k+1} \notin \bigcup_{R\in\Sigma} Z_R$). This contradicts the minimality of W. 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 4.3, 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 A.5, a d'-sequence for $\mathbf{A}$.

◀

▶ **Lemma 5.3.** *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 twin-width of $\mathbf{T}$ is at most twice the width of the ranked twin-model $(V(Y), \prec_Y, Z_{R_1}, \dots, Z_{R_k}, \tau)$.*

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 A.7. $(\widehat{Y}, \widehat{Z}_\prec, Z_{R_1}, \dots, Z_{R_k})$ 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 A.8. 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. ◀

▶ **Lemma 6.1.** *Let $\mathscr{O}$ be the class of binary tree orders and let $\mathscr{Y}^<$ be the class of rooted binary trees, with vertices ordered by some preordering. Then there exist 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

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

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

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 *preordering* 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:

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

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 7.1.** *Let $c \in \mathbb{N}$ and let $\mathscr{G}^<$ be a class of ordered graphs with star chromatic number at most c. There exist a copying-transduction T_1 with $\mathrm{bf}(\mathsf{T}_1) = c + 1$, 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 $\mathscr{G}^<$ and $\mathscr{P}$.*

Proof. Let $\mathscr{G}$ be the reduct of $\mathscr{G}^<$ obtained by "forgetting" the linear order. Let $G \in \mathscr{G}$, let $\gamma : V(G) \to [c]$ be a star coloring of G, and let $\vec{G}$ be an orientation of G obtained by orienting all bicolored stars from their roots. We mark a vertex $v \in V(G)$ by M_i if $\gamma(v) = i$ and, for $I \subseteq [c]$, by N_I if I is the set of the γ-colors of the in-neighbors of v. Let C_{c+1} be a $(c+1)$-blowing. The vertices of $\mathsf{C}_{c+1}(G)$ are the pairs $(u,i) \in V(G) \times [c+1]$, and there are new predicates P_i (with $i \in [c+1]$), where $P_i(x)$ holds if x is of the form (u,i), for some $u \in V(G)$. We define $\rho(x) := P_{c+1}(x) \vee \bigvee_{I \subseteq [c]} \bigvee_{i \in I}(N_I(x) \wedge M_i(x))$. Hence $W = \rho(\mathsf{C}_{c+1}(G))$ is the union of $V(G) \times \{c+1\}$ and the set of all pairs $(u,i) \in V(G) \times [c]$ such that u has an in-neighbor in $\vec{G}$ with color i. We consider the subgraph H of $\mathsf{C}_{c+1}(G)$ induced by W. For $x \in V(H)$ we define $f_{c+1}(x)$ as x if $P_{c+1}(x)$ or as the (only) neighbor y of x with $P_{c+1}(y)$. If x is of the form (u,i) then $f_{c+1}(x)$ is $(u, c+1)$. (Note that f_{c+1} is first-order definable.) Then, for $i \in I \subseteq [c]$ and $M_I(x)$ we define $f_i(x)$ as the (only) neighbor y of $f_{c+1}(x)$ with color i. We define the linear order $<_1$ by $x <_1 y$ if $f_{c+1}(x) < f_{c+1}(y)$ or $x = y$, $P_i(x)$, $P_j(y)$, and $i < j$ and the linear order $<_2$ by $x <_2 y$ if $P_i(x)$, $P_j(y)$ and either $f_i(x) < f_j(y)$, or $f_i(x) = f_j(y)$ and $i < j$, or $f_i(x) = f_j(y), i = j$, and $f_{c+1}(x) < f_{c+1}(y)$ (See Figure 7). We call the obtained permutation $\sigma(G^<)$. Note that this permutation depends on some arbitrary choices. We further define $\mathscr{P} = \{\sigma(G^<) : G^< \in \mathscr{G}^<\}$. The transduction T_1 is defined as the composition of C_{c+1}, the interpretation reducing the domain to the vertices satisfying $\rho(x)$, then the interpretation defining $<_1$, $<_2$ and forgetting all the other relations. Hence $\sigma(G^<) \in \mathsf{T}_1(G^<)$.

The definition of T_2 is as follows: we consider a predicate M in such a way that the maximum element of $<_1$ is in M. The domain V of $\mathsf{T}_2(\sigma)$ is $M(\sigma)$. The linear order $<$ is the restriction of $<_1$ to V. Then, x is adjacent to y if there exists $z \notin V$ and $(i,j) \in \{(1,2),(2,1)\}$ with $z <_i x$, $z <_j y$, and no vertex in V is between z and x in $<_i$ and no vertex in V between z and y in $<_j$. It is easily checked that for every $G^< \in \mathscr{G}^<$ we have $G^< \in \mathsf{T}_2(\sigma(G))$ (see Figure 7). According to Lemma 2.1 it follows that $(\mathsf{T}_2, \mathsf{T}_1)$ is a transduction pairing of $(\mathscr{C}, \mathscr{P})$. ◀