TWIN-WIDTH IV: LOW COMPLEXITY MATRICES

ÉDOUARD BONNET, UGO GIOCANTI, PATRICE OSSONA DE MENDEZ, AND STÉPHAN THOMASSÉ

ABSTRACT. We establish a list of characterizations of bounded twin-width for hereditary, totally ordered binary structures. This has several consequences. First, it allows us to show that a (hereditary) class of matrices on a finite alphabet either contains at least \( n! \) matrices of size \( n \times n \), or at most \( c^n \) for some constant \( c \). This generalizes the celebrated Stanley-Wilf conjecture/Marcus-Tardos theorem from permutation classes to any matrix class on a finite alphabet, answers our small conjecture [SODA ’21] in the case of ordered graphs, and with more work, settles a question first asked by Balogh, Bollobás, and Morris [Eur. J. Comb. ’06] on the growth of hereditary classes of ordered graphs. Second, it gives a fixed-parameter approximation algorithm for twin-width on ordered graphs. Third, it yields a full classification of fixed-parameter tractable first-order model checking on hereditary classes of ordered binary structures. Fourth, it provides an alternative proof to a model-theoretic characterization of classes with bounded twin-width announced by Simon and Toruńczyk.

1. Introduction

Matrices constitute a very common representation of a set of numbers, from linear algebra and graph theory to computer graphics and economics. Matrices can be considered in three different ways, that we will call unordered, symmetrically-reorderable, and ordered, where the row and column orders are increasingly critical.

In linear algebra, when representing linear transformations from a vector space \( \mathbb{F}^n \) to another vector space \( \mathbb{F}^m \), the order of the rows and columns is usually irrelevant, the matrix being defined up to a change of basis in the domain and the image vector spaces. Similarly when solving linear equations and inequalities, the exact order of the constraints and the naming of the variables, subject to row and column permutations, obviously do not change the set of solutions. The rank is a central complexity measure in that context.

It may happen instead that only the order of the basis can be changed, as it is the case when a matrix encodes an endomorphism, the adjacency relation of a graph or a relational structure, or is the table of a binary operation in an algebraic structure. It is then legitimate to require that the row and the column orderings are chosen consistently, so that the diagonal corresponds to pairs of the same element.

Finally, in some other contexts, the order of the rows and columns should not be touched, for example to get a well-defined matrix multiplication, because the considered basis comes with a natural total order (e.g., the basis \( (X^k)_{k \in \mathbb{N}} \) of polynomials), because the matrix encodes some geometric
object (e.g., in image representation), or because one is interested in the existence of patterns (e.g., the study of pattern-avoiding permutations).

Twin-width is a recently introduced invariant that measures how well a binary structure may be approximated by iterated lexicographic products (or replications) of basic pieces \([6,5]\). In the first paper of the series \([6]\), twin-width was defined on graphs and extended to the first two “kinds” of matrices. On unordered (possibly rectangular) matrices, it matches the twin-width of bipartite graphs where two unary relations fix the two sides of the bipartition. On symmetrically-reorderable square matrices, this corresponds to the twin-width of directed graphs (or undirected graphs, if the matrix is itself symmetric). The starting point of the current paper is to bring twin-width to ordered matrices.\(^1\) Equivalently we consider bipartite graphs where both sides of the bipartition is totally ordered, or ordered graphs (in the symmetric setting).

A second important aspect is the definition of the set (or structure) to which the entries belong. It can be a field \(\mathbb{F}\) (linear algebra), a set (relational structures), or an index set, when rows, columns, and entries refer to the same indexed set (algebraic structures). Here it will be convenient to consider that the entries belong to a finite field (as it allows to define a notion of rank), and the presentation will focus on the special case when \(\mathbb{F} = \mathbb{F}_2\). Even though we consider this special case, and a related representation by means of graphs, the results readily extend to general finite fields (or finite sets).

We now give a bit of vocabulary so that we can state, at least informally, our results. Some concepts, mainly twin-width and first-order transductions, are lengthier to explain and we will therefore postpone their definitions to the next section.

A matrix \(M\) will be indexed by two totally ordered sets, say, \(I_R\) and \(I_C\). Throughout the paper, we often observe a correspondence between \(0,1\)-matrices \(M = (m_{i,j})_{i,j}\) and ordered bipartite graphs \((I_R, I_C, E)\), where \(i \in I_R\) is adjacent to \(j \in I_C\) whenever \(m_{i,j} = 1\). (If entries can take more than two values, we may either consider a binary relational structure \((I_R, I_C, E_1, \ldots, E_s)\) or an edge coloring of \((I_R, I_C, E)\).) An \(\mathbb{F}\)-matrix has all its entries in \(\mathbb{F}\), and \(\mathcal{M}_\mathbb{F}\) denotes the set of all \(\mathbb{F}\)-matrices. Many notions related to twin-width (such as grid and mixed minor \([6]\), and in the current paper, grid rank and rich division) involve divisions of matrices. A division \(D\) of \(M\) is a pair \((D^R, D^C)\) of partitions of \(I_R\) and \(I_C\) into intervals. A division induces a representation of \(M\) as a block matrix \(M = (B_{i,j})_{1 \leq i \leq |D^R|, 1 \leq j \leq |D^C|}\), where the blocks \(B_{i,j}\) are referred to as the zones or cells of the division. A \(k\)-division is a division \(D\) such that \(|D^R| = |D^C| = k\). A \(k\)-division in which every zone has rank at least \(k\) is called a rank-\(k\) division. The growth (or speed) of a class of matrices \(\mathcal{M}\) is the function \(n \mapsto |\mathcal{M}_n|\) which counts

---

\(^1\)We postpone the exact definition of twin-width to the next section.

\(^2\)Admittedly, twin-width was already defined for binary structures in general (so for ordered matrices in particular), but we will see how a total order relation drastically helps our understanding of bounded twin-width classes.

\(^3\)Henceforth, ordered matrices will simply be called matrices.
the number of \( n \times n \) matrices of \( \mathcal{M} \). We may call \( \mathcal{M}_n \) the \( n \)-slice of class \( \mathcal{M} \).

An upper bound in twin-width, by say \( d \), is given by so-called \( d \)-sequences, iteratively identifying elements not differing too much on the relations of the binary structure. A first-order (FO) transduction of a class \( \mathcal{M} \) is any class \( \mathcal{M}' \) that can be built by non-deterministically augmenting \( \mathcal{M} \) with a constant number of unary relations and reinterpreting the relations of \( \mathcal{M} \) with first-order formulas involving these new unary relations and the old relations of \( \mathcal{M} \). An FO-interpretation is a transduction that does not use any extra unary relation. FO matrix model checking, or equivalently, FO-model checking for totally ordered binary structures, consists of determining if a given sentence is satisfied in a given binary structure, a binary relation of which being interpreted as a total order. These concepts will be properly defined in due time.

We show the following list of equivalences.

**Theorem 1.1** (informal). Given a class \( \mathcal{M} \) of matrices, the following are equivalent:

- \( i \) \( \mathcal{M} \) has bounded twin-width.
- \( ii \) (linear algebra) No matrix of \( \mathcal{M} \) has a rank-\( k \) division, for some \( k \).
- \( iii \) (Ramsey theory) \( \mathcal{M} \) does not include any of a list of families, all \( n \)-slices of which injectively map to the set of all \( n \)-permutations.
- \( iv \) (model theory) \( \mathcal{M}_{all} \) is not a first-order interpretation of \( \mathcal{M} \).
- \( v \) (model theory) \( \mathcal{M}_{all} \) is not a first-order transduction of \( \mathcal{M} \).
- \( vi \) (enumerative combinatorics) \( \mathcal{M} \) has growth smaller than \( n! \).
- \( vii \) (enumerative combinatorics) \( \mathcal{M} \) has growth \( 2^{O(n)} \).
- \( viii \) (computational complexity) FO matrix model checking is polynomial-time solvable for matrices restricted to \( \mathcal{M} \) and sentences of constant size.4

As a consequence or by-product of Theorem 1.1, we settle a handful of questions in combinatorics and algorithmic graph theory. The main by-product is an approximation algorithm for twin-width in totally ordered binary structures.

**Theorem 1.2.** There is a fixed-parameter algorithm that, given a totally ordered binary structure of twin-width \( k \), outputs a \( 2^{O(k^4)} \)-sequence.

We now detail the consequences of Theorem 1.1

1.1. **Speed gap on hereditary classes of ordered graphs.** About fifteen years ago, Balogh, Bollobás, and Morris [3, 2] analyzed the growth of ordered structures, and more specifically, ordered graphs. They conjectured [3] Conjecture 2 that a hereditary class of (totally) ordered graphs has, up to isomorphism, either at most \( O(1)^n \) \( n \)-vertex members or at least \( n^{n/2+o(n)} \), and proved it for weakly sparse graph classes, that is, without arbitrarily large bicliques (as subgraphs). In a concurrent work, Klazar [25] repeated that question, and more recently, Gunby and Pálvölgyi [21] observe that

\[ \text{The fact that this item implies the previous items is only conditional on the widely believed complexity-theoretic assumption FPT \neq AW[*].} \]
the first superexponential jump in the growth of hereditary ordered graph classes is still open.

The implication Item vi ⇒ Item vii of Theorem 1.1 settles that one-and-a-half-decade-old question. Let \( C \) be any hereditary ordered graph class with growth larger than \( c^n \), for every \( c \). We define the matrix class \( M \) as all the submatrices of the adjacency matrices of the graphs in \( C \) along the total order. We observe that for every \( c \), there is an \( n \) such that \( |M_n| > c^n \). This is because every (full) adjacency matrix of a distinct (up to isomorphism) ordered graph of \( C \) counts for a distinct matrix of \( M \). Indeed, the only automorphism of an ordered graph is the identity, due to the total order. Thus, by Theorem 1.1, \( M \) has growth at least \( n! \), asymptotically. Recall that the growth of a matrix class only accounts for its square matrices.

We now exhibit a mapping from \( M_n \) to \( \bigcup_{n \leq i \leq 2n} C_i \), where every element in the image has relatively few preimages. Let \( M \) be in \( M_n \), and let \( G_M \) be a smallest graph of \( M \) responsible for the membership \( M \in M \). The rows of \( M \) are then indexed by \( A \subseteq V(G_M) \), and its columns, by \( B \subseteq V(G_M) \), with \( V(G_M) = A \cup B \), and \( A \cap B \) potentially non-empty. \( G_M \) is a graph on at least \( n \) vertices, and at most \( 2n \). Let \( \text{Adj}(G_M) \) be its adjacency matrix where rows and columns are ordered by the total order on its vertex set. \( \text{Adj}(G_M) \) contains at most \((2n)^n \cdot (2n)^n \leq 16^n\) submatrices in \( M_n \). Therefore the same graph \( G_M \) can occur for at most \( 16^n \) matrices of \( M_n \). So \( |\bigcup_{n \leq i \leq 2n} C_i| \geq \frac{n!}{16^n} \), and \( |C_n| \geq \frac{1}{2}! \cdot (4^n \frac{n}{2})^{-1} = n^{n/2 + o(n)} \).

We will actually show the sharper bound \( |C_n| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n!}{(2k)!} \), as conjectured by Balogh et al.

1.2. Approximation of the twin-width of matrices. In the first and third paper of the series [6, 4], efficient\(^5\) algorithms are presented on graph classes of bounded twin-width. However these algorithms require a witness of bounded twin-width called \( d \)-sequences (see Section 2 for a definition). If the first two papers [6, 5] show how to find in polynomial time \( O(1) \)-sequences for a variety of bounded twin-width classes, including proper minor-closed classes, bounded rank-width classes, posets of bounded width, and long subdivisions, such an algorithm is still missing in the general case of all the graphs with twin-width at most a given threshold. As a by-product of Theorem 1.1 we obtain in Section 3 the desired missing link for ordered graphs (or matrices), that is, a fixed-parameter algorithm which either concludes that the twin-width is at least \( k \) or reports an \( f(k) \)-sequence, for some computable function \( f \). This is interesting on its own and gives some hope for the unordered case.

1.3. Fixed-parameter tractable first-order model checking. In the first-order (\( FO \)) model checking problem, one is given a structure \( G \) on a finite universe \( U \), a sentence \( \varphi \) of quantifier-depth \( \ell \), and is asked to decide if \( G \models \varphi \) holds. The brute-force algorithm takes time \( |U|^{O(\ell)} \), by exploring the full game tree. The question is whether a uniformly polynomial-time algorithm exists, that is, with running time \( f(\ell)|U|^{O(1)} \). In the language of

\(^5\)provably more efficient than what is possible on general graphs, under standard complexity-theoretic assumptions
parameterized complexity, a parameterized problem is called fixed-parameter tractable (FPT) if there exists an algorithm A (called a fixed-parameter algorithm), a computable function $f : \mathbb{N} \to \mathbb{N}$, and a constant $c$ such that, given an input of size $n$ and parameter $k$, the algorithm A correctly decides if the inputs has the desired property in time bounded by $f(k)n^c$. The complexity class containing all fixed-parameter tractable problems is called FPT. (We refer the interested to [8] for more details on parameterized algorithms.)

When the input structures range over the set of all finite graphs, FO-model checking is known to be $\text{AW}[\ast]$-complete [10], thus not FPT unless the widely-believed complexity-theoretic assumption $\text{FPT} \neq \text{AW}[\ast]$ fails.

There is an ongoing program aiming to classify all the hereditary graph classes on which FO-model checking is FPT. Currently such an algorithm is known for nowhere dense classes [20], for structurally bounded-degree classes [16] (and more generally for perturbations of degenerate nowhere dense classes [17]), for map graphs [12], for some families of intersection and visibility graphs [22], for transductions of bounded expansion classes when a depth-2 low shrub-depth cover of the graph is given [18], and for classes \(^6\) with bounded twin-width [6]. It is believed that every class which is, in that context, “essentially different” from the class of all graphs \(^7\) admits a fixed-parameter tractable FO-model checking. Settling this conjecture might require to get a unified understanding of bounded twin-width and structurally nowhere dense classes.

Much effort [15, 13, 11, 26, 31] has also been made in graph classes augmented by an order or a successor relation. We refer the interested reader to the joint journal version [14], subsuming all five previous references. There are two different settings: the general ordered case (with no restriction), and the order invariant case (where the queried formulas may use the new relation but must not depend on the particular ordering). In the order-invariant setting, the model checking is shown fixed-parameter tractable on classes of bounded expansion and colored posets of bounded width [14]. In the general ordered case, the same authors observe that $\text{FO}[<]$-model checking is $\text{AW}[\ast]$-complete when the underlying graph class is as simple as partial matchings [14, Theorem 1]. By considering the edge and order relations as a whole unit, fixed-parameter tractable algorithms do exist in a relatively broad scenario, namely, when the resulting binary structures have bounded twin-width. The equivalence between Item i and Item viii, and the fact that $O(1)$-sequences can be efficiently computed (see Section 1.2), completely resolves this version of the general ordered case.

1.4. **Bounded twin-width classes are exactly those than can be totally ordered and remain monadically NIP.** We refer the reader to Section 2.3 for the relevant background. Simon and Toruńczyk [30] recently announced the following characterization of bounded twin-width classes: A class $C$ of binary structures over a signature $\sigma$ has bounded twin-width if and only if there exists a monadically dependent (i.e., monadically NIP) class $D$ over $\sigma \cup \{<\}$, where $<$ is interpreted as a total order, such

\(^6\)With the caveat that a witness of low twin-width is needed (see Section 1.2).

\(^7\)Precisely, every class without an FO-transduction which is the class of all graphs (see Section 2 for the relevant definitions)
that $C = \text{Reduct}_\sigma(D)$, where $\text{Reduct}_\sigma(\cdot)$ simply forgets the relation $<$. The forward implication can be readily derived from known results [6]. For any binary structure, there is a total order on its vertices which, added to the structure, does not change its twin-width. This is by definition (see Section 2.1). Now every class of bounded twin-width is monadically NIP. This is because FO-transductions preserve bounded twin-width. The implication Item v $\implies$ Item i yields the backward direction, since a rephrasing of Item v is that the class is monadically NIP. Thus we also obtain Simon and Törnäczyk’s characterization.

1.5. Small conjecture. Classes of bounded twin-width are small [5], that is, they contain at most $n!c^n$ distinct labeled $n$-vertex structures, for some constant $c$. In the same paper, the converse is conjectured for hereditary classes. In the context of classes of totally ordered structures, it is simpler to drop the labeling and to count up to isomorphism. Indeed every structure has no non-trivial automorphism. Then a class is said small if, up to isomorphism, it contains at most $c^n$ distinct $n$-vertex structure. With that in mind, the equivalence between Item i and Item vii resolves the conjecture in the particular case of ordered graphs (or matrices).

2. Preliminaries

Everything which is relevant to the rest of the paper will now be properly defined. We may denote by $[i,j]$ the set of integers that are at least $i$ and at most $j$, and $[i]$ is a short-hand for $[1,i]$. We start with the definition of twin-width.

2.1. Twin-width. In the first paper of the series [6], we define twin-width for general binary structures. The twin-width of (ordered) matrices can be defined by encoding the total orders on the rows and on the columns with two binary relations. However we will give an equivalent definition, tailored to ordered structures. This slight shift is already a first step in understanding these structures better, with respect to twin-width.

Let $M$ be a $n \times m$ matrix with entries ranging in a fixed finite set. We denote by $R := \{r_1, \ldots, r_n\}$ its set of rows and by $C := \{c_1, \ldots, c_m\}$ its set of columns. Let $S$ be a non-empty subset of columns, $c_a$ be the column of $S$ with minimum index $a$, and $c_b$, the column of $S$ with maximum index $b$. The span of $S$ is the set of columns $\{c_a, c_{a+1}, \ldots, c_{b-1}, c_b\}$. We say that a subset $S \subseteq C$ is in conflict with another subset $S' \subseteq C$ if their spans intersect. A partition $\mathcal{P}$ of $C$ is a $k$-overlapping partition if every part of $\mathcal{P}$ is in conflict with at most $k$ other parts of $\mathcal{P}$. The definitions of span, conflict, and $k$-overlapping partition similarly apply to sets of rows. With that terminology, a division is a 0-overlapping partition.

A partition $\mathcal{P}$ is a contraction of a partition $\mathcal{P}'$ (defined on the same set) if it is obtained by merging two parts of $\mathcal{P}'$. A contraction sequence of $M$ is a sequence of partitions $\mathcal{P}_1, \ldots, \mathcal{P}_{n+m-1}$ of the set $R \cup C$ such that $\mathcal{P}_1$ is the partition into $n + m$ singletons, $\mathcal{P}_{i+1}$ is a contraction of $\mathcal{P}_i$ for all $i \in [n+m-2]$, and such that $\mathcal{P}_{n+m-1} = \{R, C\}$. In other words, we merge at every step two column parts (made exclusively or columns) or two row parts (made exclusively or rows), and terminate when all rows and all columns
both form a single part. We denote by $\mathcal{P}^R_i$ the partition of $R$ induced by $\mathcal{P}_i$ and by $\mathcal{P}^C_i$ the partition of $C$ induced by $\mathcal{P}_i$. A contraction sequence is $k$-overlapping if all partitions $\mathcal{P}^R_i$ and $\mathcal{P}^C_i$ are $k$-overlapping partitions. Note that a 0-overlapping sequence is a sequence of divisions.

If $S^R$ is a subset of $R$, and $S^C$ is a subset of $C$, we denote by $S^R \cap S^C$ the submatrix at the intersection of the rows of $S^R$ and of the columns of $S^C$. Given some column part $C_a$ of $\mathcal{P}^C_i$, the error value of $C_a$ is the number of row parts $R_b$ of $\mathcal{P}^R_i$ for which the submatrix $C_a \cap R_b$ of $M$ is not constant. The error value is defined similarly for rows, by switching the role of columns and rows. The error value of $\mathcal{P}_i$ is the maximum error value of some part in $\mathcal{P}^R_i$ or in $\mathcal{P}^C_i$. A contraction sequence is a $(k, e)$-sequence if all partitions $\mathcal{P}^R_i$ and $\mathcal{P}^C_i$ are $k$-overlapping partitions with error value at most $e$. Strictly speaking, to be consistent with the definitions in the first paper [6], the twin-width of a matrix $M$, denoted by $\text{tww}(M)$, is the minimum $k + e$ such that $M$ has a $(k, e)$-sequence. This matches, setting $d := k + e$, what we called a $d$-sequence for the binary structure encoding $M$. We will however not worry about the exact value of twin-width. Thus for the sake of simplicity, we often consider the minimum integer $k$ such that $M$ has a $(k, k)$-sequence. This integer is indeed sandwiched between $\text{tww}(M)/2$ and $\text{tww}(M)$.

The twin-width of a matrix class $\mathcal{M}$, denoted by $\text{tww}(\mathcal{M})$, is simply defined as the supremum of $\{\text{tww}(M) \mid M \in \mathcal{M}\}$. We say that $\mathcal{M}$ has bounded twin-width if $\text{tww}(\mathcal{M}) < \infty$, or equivalently, if there is a finite integer $k$ such that every matrix $M \in \mathcal{M}$ has twin-width at most $k$. A class $\mathcal{G}$ of ordered graphs has bounded twin-width if all the adjacency matrices of graphs $G \in \mathcal{G}$ along their vertex ordering, or equivalently their submatrix closure, form a set/class with bounded twin-width.

### 2.2. Rank division and rich division

We will now require that the matrix entries are elements of a finite field $\mathbb{F}$. We recall that a division $\mathcal{D}$ of a matrix $M$ is a pair $(\mathcal{D}^R, \mathcal{D}^C)$, where $\mathcal{D}^R$ (resp. $\mathcal{D}^C$) is a partition of the rows (resp. columns) of $M$ into (contiguous) intervals, or equivalently, a 0-overlapping partition. A $d$-division is a division satisfying $|\mathcal{D}^R| = |\mathcal{D}^C| = d$. For every pair $R_i \in \mathcal{D}^R$, $C_j \in \mathcal{D}^C$, the submatrix $R_i \cap C_j$ may be called zone (or cell) of $\mathcal{D}$ since it is, by definition, a contiguous submatrix of $M$. We observe that a $d$-division has $d^2$ zones.

A rank-$k$ $d$-division of $M$ is a $d$-division $\mathcal{D}$ such that for every $R_i \in \mathcal{D}^R$ and $C_j \in \mathcal{D}^C$ the zone $R_i \cap C_j$ has rank at least $k$ (over $\mathbb{F}$). A rank-$k$ division is simply a short-hand for a rank-$k$ $d$-division. The grid rank of a matrix $M$, denoted by $\text{gr}(M)$, is the largest integer $k$ such that $M$ admits a rank-$k$ division. The grid rank of a matrix class $\mathcal{M}$, denoted by $\text{gr}(\mathcal{M})$, is defined as $\sup\{\text{tww}(M) \mid M \in \mathcal{M}\}$. A class $\mathcal{M}$ has bounded grid rank if $\text{gr}(\mathcal{M}) < \infty$, or equivalently, if there exists an integer $k$ such that for every matrix $M \in \mathcal{M}$, and for every $k$-division $\mathcal{D}$ of $M$, there is a zone of $\mathcal{D}$ with rank less than $k$.

Closely related to rank divisions, a $k$-rich division is a division $\mathcal{D}$ of a matrix $M$ on rows and columns $R \cup C$ such that:

- for every part $R_a$ of $\mathcal{D}^R$ and for every subset $Y$ of at most $k$ parts in $\mathcal{D}^C$, the submatrix $R_a \cap (C \setminus Y)$ has at least $k$ distinct row vectors, and symmetrically
• for every part \( C_b \) of \( D^C \) and for every subset \( X \) of at most \( k \) parts in \( D^R \), the submatrix \((R \setminus \cup X) \cap C_b\) has at least \( k \) distinct column vectors.

Informally, in a large rich division (that is, a \( k \)-rich division for some large value of \( k \)), the diversity in the column vectors within a column part cannot drop too much by removing a controlled number of row parts. And the same applies to the diversity in the row vectors.

We now move on to describe the relevant concepts in finite model theory.

2.3. Model Theory. A relational signature \( \sigma \) is a set of relation symbols \( R_i \) with associated arities \( r_i \). A \( \sigma \)-structure \( A \) is defined by a finite set \( A \) (the domain of \( A \)) together with a subset \( R^A_i \) of \( A^{r_i} \) for each relation symbol \( R_i \in \sigma \) with arity \( r_i \). The first-order language \( \text{FO}(\sigma) \) associated to \( \sigma \)-structures defines, for each relation symbol \( R_i \) with arity \( r_i \) the predicate \( R_i \) such that \( A \models R_i(v_1, \ldots, v_{r_i}) \) if \((v_1, \ldots, v_{r_i}) \in R^A_i \).

Let \( \varphi(\overline{x}, \overline{y}) \) be a first-order formula in \( \text{FO}(\sigma) \) and let \( \mathcal{C} \) be a class of \( \sigma \)-structures. The formula \( \varphi \) is independent over \( \mathcal{C} \) if, for every integer \( k \in \mathbb{N} \) there exist a \( \sigma \)-structure \( A \in \mathcal{C} \), \( k \) tuples \( \overline{v}_1, \ldots, \overline{v}_k \in A^{[k]} \), and \( 2^k \) tuples \( \overline{v}_0, \ldots, \overline{v}_{[k]} \in A^{[k]} \) with

\[
A \models \varphi(\overline{v}_i, \overline{v}_j) \iff i \in I.
\]

The class \( \mathcal{C} \) is independent if there is a formula \( \varphi(\overline{x}, \overline{y}) \in \text{FO}(\sigma) \) that is independent over \( \mathcal{C} \). Otherwise, the class \( \mathcal{C} \) is dependent (or \( \text{NIP} \), for Not the Independence Property).

A theory \( T \) is a consistent set of first-order sentences. We will frequently consider classes of structures satisfying some theory. For instance, a (simple undirected) graph is a structure on the signature \( \sigma_{\text{graph}} \) with unique binary relation symbol \( E \) satisfying the theory \( T_{\text{graph}} \) consisting of the two sentences \( \forall x \neg E(x, x) \) (which asserts that a graph has no loops) and \( \forall x \forall y (E(x, y) \leftrightarrow E(y, x)) \) (which asserts that the adjacency relation of a graph is symmetric). We now define the signatures and theories corresponding to 0, 1-matrices and to ordered graphs. A linear order is a \( \sigma_{<} \)-structure satisfying the theory \( T_{<} \), where \( \sigma_{<} \) consists of the binary relation \( < \), and \( T_{<} \) consists of the following sentences, which express that \( < \) is a linear order.

\[
\forall x \neg(x < x);
\forall x \forall y \ (x = y) \lor (x < y) \lor (y < x);
\forall x \forall y \forall z \ ((x < y) \land (y < z)) \rightarrow (x < z).
\]

A 0, 1-matrix is a \( \sigma_{\text{matrix}} \)-structure satisfying the theory \( T_{\text{matrix}} \), where \( \sigma_{\text{matrix}} \) consists of a unary relational symbol \( R \) (interpreted as the indicator of row indices), a binary relation \( < \) (interpreted as a linear order), and a binary relation \( M \) (interpreted as the matrix entries), and the theory \( T_{\text{matrix}} \) is obtained by adding to \( T_{<} \) the sentences

\[
\forall x \forall y \ (R(x) \land \neg R(y)) \rightarrow (x < y),
\forall x \forall y \ M(x, y) \rightarrow (R(x) \land \neg R(y)).
\]

The first sentence asserts that all the row indices are before (along \( < \)) all the column indices. The second sentence asserts that the first variable of \( M \) is a row index, while the second variable of \( M \) is a column index.
An ordered graph is a $\sigma_{\text{graph}}$-structure satisfying the theory $T_{\text{graph}}$, where $\sigma_{\text{graph}}$ consists of the binary relations $<$ and $E$, and where $T_{\text{graph}}$ consists of the union of $T_{\text{graph}}$ and $T_{\prec}$.

Let $\sigma_1, \sigma_2$ be signatures and let $T_1, T_2$ be theories, in FO($\sigma_1$) and FO($\sigma_2$), respectively. A simple interpretation of $\sigma_2$-structures in $\sigma_1$-structures is a tuple $1 = (\nu, \rho_1, \ldots, \rho_k)$ of formulas in FO($\sigma_1$), where $\nu(x)$ is a single free variable and, for each relation symbol $R_i \in \sigma_2$ with arity $r_i$ the formula $\rho_i$ has $r_i$ free variables. If $A$ is a $\sigma_1$-structure, the $\sigma_2$-structure $I(A)$ has domain $\nu(A) = \{ v \in A : A \models \nu(v) \}$ and relation $R_i^{I(A)} = \rho_i(A) \cap \nu(A)^{r_i}$, that is:

$$R_i^{I(A)} = \{ (v_1, \ldots, v_{r_i}) \in \nu(A)^{r_i} : A \models \rho_i(v_1, \ldots, v_{k}) \}.$$ 

An important property of (simple) interpretations is that, for every formula $\varphi(x_1, \ldots, x_k) \in \text{FO}(\sigma_2)$ there is a formula $I^*(\varphi)(x_1, \ldots, x_k)$ such that for every $\sigma_1$-structure $A$ and every $x_1, \ldots, x_k \in \nu(A)$ we have

$$I(A) \models \varphi(x_1, \ldots, x_k) \iff A \models I^*(\varphi)(x_1, \ldots, x_k).$$

We say that $I$ is a simple interpretation of $\sigma_2$-structures satisfying $T_2$ in $\sigma_1$-structures satisfying $T_1$ if, for every $\theta \in T_2$ we have $T_1 \vdash I^*(\theta)$. Then, for every $\sigma_1$-structure $A$ we have

$$A \models T_1 \implies I(A) \models T_2.$$ 

By extension we say, for instance, that $I$ is a simple interpretation of ordered graphs in 0, 1-matrices if it is a simple interpretation of $\sigma_{\text{graph}}$-structures satisfying $T_{\text{graph}}$ in $\sigma_{\text{matrix}}$-structures satisfying $T_{\text{matrix}}$.

Let $\sigma_2 \subseteq \sigma_1$ be relational signatures. The $\sigma_2$-reduct (or $\sigma_2$-shadow) of a $\sigma_1$-structure $A$ is the structure obtained from $A$ by “forgetting” all the relations not in $\sigma_2$. This interpretation of $\sigma_2$-structures in $\sigma_1$-structures is denoted by $\text{Reduct}_{\sigma_2}$ or simply $\text{Reduct}$, when $\sigma_2$ is clear from context.

A monadic lift of a class $\mathcal{C}$ of $\sigma$-structures is a class $\mathcal{C}^+$ of $\sigma^+$-structures, where $\sigma^+$ is the union of $\sigma$ and a set of unary relation symbols, and $\mathcal{C} = \{ \text{Reduct}_{\sigma}(A) : A \in \mathcal{C}^+ \}$. A class $\mathcal{C}$ of $\sigma$-structures is monadically dependent (or monadically NIP) if every monadic lift of $\mathcal{C}$ is dependent (or NIP). A transduction $T$ from $\sigma_1$-structures to $\sigma_2$-structures is defined by an interpretation $I_T$ of $\sigma_2$-structures in $\sigma_1^+$-structures, where $\sigma_1^+$ is the union of $\sigma_1^+$ and a set of unary relation symbols. For a class $\mathcal{C}$ of $\sigma_1$-structures, we define $T(\mathcal{C})$ as the class $I_T(\mathcal{C}^+)$ where $\mathcal{C}^+$ is the set of all $\sigma_2^+$-structures $A^+$ with $\text{Reduct}_{\sigma_1}(A^+) \in \mathcal{C}$. A class $\mathcal{D}$ of $\sigma_2$-structures is a $T$-transduction of a class $\mathcal{C}$ of $\sigma_1$-structures if $\mathcal{D} \subseteq T(\mathcal{C})$. More generally, a class $\mathcal{D}$ of $\sigma_2$-structures is a transduction of a class $\mathcal{C}$ of $\sigma_1$-structures if there exists a transduction $T$ from $\sigma_1$-structures to $\sigma_2$ structures with $\mathcal{D} \subseteq T(\mathcal{C})$. Note that the composition of two transductions is also a transduction.

The following theorem witnesses that transductions are particularly fitting to the study of monadic dependence:

**Theorem 2.1** (Baldwin and Shelah [1]). A class $\mathcal{C}$ of $\sigma$-structures is monadically dependent if and only if for every monadic lift $\mathcal{C}^+$ of $\mathcal{C}$ (in $\sigma^+$-structures), every formula $\varphi(\overline{x}, \overline{y}) \in \text{FO}(\sigma^+)$ with $|\overline{x}| = |\overline{y}| = 1$ is dependent over $\mathcal{C}^+$. 

Consequently, \( C \) is monadically dependent if and only if the class \( C \) of all finite graphs is not a transduction of \( C \).

**Corollary 2.2.** If \( \mathcal{D} \) is a transduction of a class \( C \) and \( C \) is monadically dependent then \( \mathcal{D} \) is monadically dependent.

**Proof.** Otherwise, the class \( C \) of all finite graphs is a transduction of \( \mathcal{D} \) and, by composition, a transduction of \( C \), contradicting the monadic dependence of \( C \). \( \Box \)

2.4. **Enumerative Combinatorics.** In the context of unordered structures, a graph class \( C \) is said small if there is a constant \( c \), such that its number of \( n \)-vertex graphs bijectively labeled by \([n]\) is at most \( n!c^n \). When considering totally ordered structures, for which the identity is the unique automorphism, one can advantageously drop the labeling and the \( n! \) factor. Indeed, on these structures, counting up to isomorphism or up to equality is the same. Thus a matrix class \( \mathcal{M} \) is said small if there exists a real number \( c \) such that the total number of \( m \times n \) matrices in \( \mathcal{M} \) is at most \( c^{\max(m,n)} \). Analogously to permutation classes which are by default supposed closed under taking subpermutations (or patterns), we will define a class of matrices as a set of matrices closed under taking submatrices. The submatrix closure of a matrix \( M \) is the set of all submatrices of \( M \) (including \( M \) itself). Thus our matrix classes include the submatrix closure of every matrix they contain. On the contrary, classes of (ordered) graphs are only assumed to be closed under isomorphism. A hereditary class of (ordered) graphs (resp. binary structures) is one that is closed under taking induced subgraphs (resp. induced substructures).

Marcus and Tardos [27] showed the following central result, henceforth referred to as Marcus-Tardos theorem, which by an argument due to Klazar [24] was known to imply the Stanley-Wilf conjecture, that permutation classes avoiding any fixed pattern are small.

**Theorem 2.3.** There exists a function \( \text{mt} : \mathbb{N} \to \mathbb{N} \) such that every \( n \times m \) matrix \( M \) with at least \( \text{mt}(k) \max(n,m) \) nonzero entries has a \( k \)-division in which every zone contains a non-zero entry.

We call \( \text{mt}(\cdot) \) the Marcus-Tardos bound. The current best bound is \( \text{mt}(k) = \frac{8}{3} (k+1)^2 2^{4k} = 2^{O(k)} \). Among other things, The Marcus-Tardos theorem is a crucial tool in the development of the theory around twin-width. In the second paper of the series [5], we generalize the Stanley-Wilf conjecture/Marcus-Tardos theorem to classes with bounded twin-width. We show that every graph class with bounded twin-width is small (while proper subclasses of permutation graphs have bounded twin-width [6]). This can be readily extended to every bounded twin-width class of binary structures. We conjectured that the converse holds for hereditary classes: Every hereditary small class of binary structures has bounded twin-width. We will show this conjecture, in the current paper, for the special case of totally ordered binary structures.

We denote by \( \mathcal{M}_n \), the \( n \)-slice of a matrix class \( \mathcal{M} \), that is the set of all \( n \times n \) matrices of \( \mathcal{M} \). The growth (or speed) of a matrix class is the function \( n \in \mathbb{N} \mapsto |\mathcal{M}_n| \). A class \( \mathcal{M} \) has subfactorial growth if there is a finite integer
beyond which the growth of \( M \) is strictly less than \( n! \); more formally, if there is \( n_0 \) such that for every \( n \geq n_0 \), \( |M_n| \leq n! \). Similarly, \( \mathcal{C} \) being a class of ordered graphs, the \( n \)-slice of \( \mathcal{C} \), \( \mathcal{C}_n \), is the set of \( n \)-vertex ordered graphs in \( \mathcal{C} \). And the growth (or speed) of a class \( \mathcal{C} \) of ordered graphs is the function \( n \in \mathbb{N} \mapsto |\mathcal{C}_n| \).

2.5. Computational Complexity. We recall that first-order (FO) matrix model checking asks, given a matrix \( M \) (or a totally ordered binary structure \( S \)) and a first-order sentence \( \phi \) (i.e., a formula without any free variable), if \( M \models \phi \) holds. The atomic formulas in \( \phi \) are of the kinds described in Section 2.3.

We then say that a matrix class \( M \) is tractable if FO-model checking is fixed-parameter tractable (FPT) when parameterized by the sentence size and the input matrices are drawn from \( M \). That is, \( M \) is tractable if there exists a constant \( c \) and a computable function \( f \), such that \( M \models \phi \) can be decided in time \( f(\ell)(m+n)^c \), for every \( n \times m \)-matrix \( M \in M \) and FO sentence \( \phi \) of quantifier depth \( \ell \). We may denote the size of \( M \), \( n + m \), by \( |M| \), and the quantifier depth (i.e., the maximum number of nested quantifiers) of \( \phi \) by \( |\phi| \). Similarly a class \( C \) of binary structures is said tractable if FO-model checking is FPT on \( C \).

FO-model checking of general (unordered) graphs is \( \text{AW}[*] \)-complete \([10]\), and thus very unlikely to be FPT. Indeed \( \text{FPT} \neq \text{AW}[*] \) is a much weaker assumption than the already widely-believed Exponential Time Hypothesis \([23]\), and if false, would in particular imply the existence of a subexponential algorithm solving 3-SAT. In the first paper of the series \([6]\), we show that FO-model checking of general binary structures of bounded twin-width given with an \( O(1) \)-sequence can even be solved in linear FPT time \( f(|\phi|)|U| \), where \( U \) is the universe of the structure. In other words, bounded twin-width classes admitting a \( g(\text{OPT}) \)-approximation for the contraction sequences are tractable. It is known for (unordered) graph classes that the converse does not hold. For instance, the class of all subcubic graphs (i.e., graphs with degree at most 3) is tractable \([29]\) but has unbounded twin-width \([5]\). Theorem 1.2 will show that, on every class of ordered graphs, a fixed-parameter approximation algorithm for the contraction sequence exists. Thus every bounded twin-width class of ordered graphs is tractable. We will also see that the converse holds for hereditary classes of ordered graphs.

2.6. Ramsey Theory. The order type of a pair \((x, y)\) of elements of a linearly ordered set is the integer \( \text{ot}(x, y) \) defined by

\[
\text{ot}(x, y) = \begin{cases} 
-1 & \text{if } x > y \\
0 & \text{if } x = y \\
1 & \text{if } x < y.
\end{cases}
\]

A class \( M \) is pattern-avoiding if it does not include any of the matrix classes of the set
\[
\mathcal{P} := \{ F_\eta \mid \eta : \{-1, 1\} \times \{-1, 1\} \to \{0, 1\}\}
\]

of 16 classes, where \( F_\eta \) is the hereditary closure of \( \{ F_\eta(\sigma) \mid \sigma \in \mathbb{S}_n, n \geq 1 \} \). For a fixed function \( \eta : \{-1, 1\} \times \{-1, 1\} \to \{0, 1\} \), the matrix \( F_\eta(\sigma) = \)
$(f_{i,j})_{1 \leq i,j \leq n}$ corresponds to an encoding of the permutation matrix $M_\sigma$ of $\sigma \in \mathfrak{S}_n$, where $f_{i,j}$ only depends on the order types between $i$ and $\sigma^{-1}(j)$, and between $j$ and $\sigma(i)$ in a way prescribed by $\eta$. In other words, $f_{i,j}$ is fully determined by asking whether $(i,j)$ is, in $M_\sigma$, below or above the 1 of its column and whether it is to the left or the right of the 1 of its row.

We now give the formal definition of $F_\eta(\sigma) = (f_{i,j})_{1 \leq i,j \leq n}$, but we will recall it and provide some visual intuition in due time. For every $i,j \in [n]$:

$$f_{i,j} := \begin{cases} \eta(\text{ot}(\sigma^{-1}(j), i), \text{ot}(j, \sigma(i))) & \text{if } \sigma(i) \neq j \\ 1 - \eta(1, 1) & \text{if } \sigma(i) = j \end{cases}$$

We give a similar definition in Section 7 for ordered graphs: a hereditary class $C$ of ordered graphs is matching-avoiding if it does not include any ordered graph class $\mathcal{M}_{\eta,\lambda,\rho}$ of a set of 256 classes (corresponding this time to encodings of ordered matchings). The precise definition is more technical, and not that important at this stage, hence our decision of postponing it to Section 7.

2.7. Our results. We can now restate the list of equivalences announced in the introduction, with the vocabulary of this section.

**Theorem 1.1.** Given a class $\mathcal{M}$ of matrices, the following are equivalent.

1. $\mathcal{M}$ has bounded twin-width.
2. $\mathcal{M}$ has bounded grid rank.
3. $\mathcal{M}$ is pattern-avoiding.
4. $\mathcal{M}$ is dependent.
5. $\mathcal{M}$ is monadically dependent.
6. $\mathcal{M}$ has subfactorial growth.
7. $\mathcal{M}$ is small.
8. $\mathcal{M}$ is tractable. (The implication from Item (viii) holds if $\text{FPT} \neq \text{AW}[*]$.)

For the reader to get familiar with the definitions and notations, we give a compact version of Theorem 1.1. We also introduce a technical condition, Item (ix), which will be a key intermediate step in proving that Item (ii) implies Item (i) as well as in getting an approximation algorithm for the twin-width of a matrix.

**Theorem 1.1** (compact reminder of the definitions and notations + Item (ix)). Given a class $\mathcal{M}$ of matrices, the following are equivalent.

1. $\text{tww}(\mathcal{M}) < \infty$.
2. $\text{gr}(\mathcal{M}) < \infty$.
3. For every $\mathcal{F}_\eta \in \mathcal{P}$, $\exists M \in \mathcal{F}_\eta$, $M \notin \mathcal{M}$.
4. For every $\text{FO}$-interpretation $I$, $I(\mathcal{M}) \neq \mathcal{M}_{\text{all}}$.
5. For every $\text{FO}$-transduction $T$, $T(\mathcal{M}) \neq \mathcal{M}_{\text{all}}$.
6. $\exists n_0 \in \mathbb{N}$, $|\mathcal{M}_{n_0}| \leq n_0$, $\forall n \geq n_0$.
7. $\exists c \in \mathbb{N}$, $|\mathcal{M}_n| < c^n$, $\forall n \in \mathbb{N}$.
8. Given $(M \in \mathcal{M}, \phi \in \text{FO}[\tau])$, $M \models \phi$ can be decided in time $f(|\phi|)|M|$.
9. $\exists q \in \mathbb{N}$, no $M \in \mathcal{M}$ admits a $q$-rich division.

We transpose these results for hereditary classes of ordered graphs. We also refine the model-theoretic (Items 3 and 4) and growth (Item 7) characterizations.
Theorem 2.4. Let $\mathcal{C}$ be a hereditary class of ordered graphs. The following are equivalent.

1. $\mathcal{C}$ has bounded twin-width.
2. $\mathcal{C}$ is monadically dependent.
3. $\mathcal{C}$ is dependent.
4. No simple interpretation in $\mathcal{C}$ is the class of all ordered graphs.
5. $\mathcal{C}$ is small.
6. $\mathcal{C}$ contains $2^{O(n)}$ ordered $n$-vertex graphs.
7. $\mathcal{C}$ contains less than $\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!$ ordered $n$-vertex graphs, for some $n$.
8. $\mathcal{C}$ does not include one of 256 hereditary ordered graph classes $\mathcal{M}_{\eta,\lambda,\rho}$ with unbounded twin-width.
9. There exists a permutation $\sigma$ such that $\mathcal{C}$ does not include any of 256 ordered graphs defined from $\sigma$.
10. FO-model checking is fixed-parameter tractable on $\mathcal{C}$.
   
   (This implies the other items only if FPT $\neq$ AW[*].)

The previous theorem holds more generally for hereditary ordered classes of binary structures. In an informal nutshell, the high points of the paper read: For hereditary ordered binary structures, bounded twin-width, small, subfactorial growth, and tractability of FO-model checking are all equivalent. We conclude by giving a more detailed statement of the approximation algorithm.

Theorem 1.2 (more precise statement). There is a fixed-parameter algorithm, which, given an ordered binary structure $G$ and a parameter $k$, either outputs

- a $2^{O(k^4)}$-sequence of $G$, implying that $\text{tww}(G) = 2^{O(k^4)}$, or
2.8. Outline. Bounded twin-width is already known to imply interesting properties: FPT FO-model checking if the $O(1)$-sequences are part of the input [6], monadic dependence [6], smallness [9] (see the green and orange arrows in Figures 1 and 2). For a characterization of some sort in the particular case of ordered structures, the challenge is to find interesting properties implying bounded twin-width. A central characterization in the first paper of the series [6] goes as follows. Let us call 1,2-adjacency matrix a usual 0,1-adjacency matrix where the 0 entries (non-edges) are replaced, for a purely technical reason, by 2. A graph class $C$ has bounded twin-width if and only if there is a constant $d_C$ such that every graph in $C$ admits a 1,2-adjacency matrix with no rank-2 $d_C$-division. A reformulation of the latter condition is that there is an ordering of the vertex set such that the adjacency matrix has some property (no large division where every cell has rank at least 2). The backward direction is effective: From such an ordering, we obtain an $O(1)$-sequence in polynomial time.

Now that we consider ordered matrices (and our graphs come with a total order) it is tempting to try this order to get a witness of low twin-width. Things are not that simple. Consider the checkerboard matrix (with 1 entries at positions $(i,j)$ such that $i + j$ is even, and 0 otherwise). It admits a $(1,2)$-sequence. We can merge the first and third columns into $C_o$, the second and fourth columns into $C_e$, then $C_o$ and the fifth into $C_o$, $C_e$ and the sixth into $C_e$, and so on. This creates a sequence of 1-overlapping partitions since only two column parts, $C_o$ and $C_e$, ever get in conflict. The maximum error value remains 0 since all columns of odd (resp. even) index are equal. Then we proceed in the same way on the row parts. Again it makes for a “partial” $(1,0)$-sequence. Finally we are left with two row parts and two column parts that we merge in any order. This yields an error value of 2, while preserving the fact that the partitions are 1-overlapping.

So the twin-width of all the checkerboard matrices is bounded. Yet they have rank-2 $d$-divisions for arbitrarily large $d$ (by dividing after every even-indexed row and column). Now a good reordering would put all the odd-indexed columns together, followed by all the even-indexed columns. Reordered in this way, a matrix encoding both the initial matrix and the original order would have only small rank-2 $d$-divisions.

Can we find such reorderings automatically? Eventually we can but a crucial opening step is precisely to nullify the importance of the reordering. We show that matrices have bounded twin-width exactly when they do not admit rank-$k$ $k$-divisions for arbitrary $k$. This natural strengthening on the condition that cells should satisfy (rank at least $k$ instead of rank at least 2) exempts us from the need to reorder. Note that the checkerboard matrix does not have any rank-$k$ division already for $k = 3$, for the good reason that it has rank 2.

An important intermediate step is provided by the concept of rich divisions. We first prove that a greedy strategy to find a potential $O(1)$-sequence can only be stopped by the presence of a large rich division; thus, unbounded twin-width implies the existence of arbitrarily large rich divisions. This brings a theme developed in [6] to the ordered world. In turn we show that
huge rich divisions contain large rank divisions. As often in the series, this leverages Marcus-Tardos theorem and is entirely summarized by Figure 4.

By a series of Ramsey-like arguments, we find in large rank divisions more and more structured submatrices encoding universal permutations. Eventually we find at least one of sixteen encodings of all permutations (i.e., $\mathcal{F}_\eta$ for one of the sixteen “$\eta$”). More precisely, the encoding of each $n$-permutation is contained in $(\mathcal{F}_\eta)_n$, the $n \times n$ matrices of $\mathcal{F}_\eta$.

This chain of implications shows that hereditary classes with unbounded twin-width have growth at least $n!$. Conversely it was known that labeled classes with growth $n! 2^{o(n)}$ have unbounded twin-width [3], thus (unlabeled) ordered classes with growth $2^{o(n)}$ also have unbounded twin-width. That establishes the announced speed gap for ordered hereditary classes of binary structures.

Finally we translate the permutation encodings in the language of ordered graphs. This allows us to refine the growth gap specifically for ordered graphs. We also prove that including a family $\mathcal{F}_\eta$ or its ordered-graph equivalent is an obstruction to being NIP. This follows from the fact that the class of all permutation graphs is independent. As we get an effectively constructible transduction to the set of all structures (matrices or ordered graphs), we conclude that FO-model checking is not FPT on hereditary classes of unbounded twin-width. This is the end of the road. The remaining implications to establish the equivalences of Theorems 1.1 and 2.4 come from [6, Sections 7 and 8], [5, Section 3], and Theorem 1.2 (see Figure 2).

**Figure 2.** A more detailed proof diagram.
2.9. Organization. The rest of the paper is organized as follows. In Section 2, we show that Item i and Item ix are equivalent. As a by-product, we obtain a fixed-parameter \( f(\text{OPT}) \)-approximation algorithm for the twin-width of ordered matrices. In Section 3, we prove the implication Item ii \( \Rightarrow \) Item ix. In Section 4, we introduce the rank Latin divisions and show that large rank divisions contain large rank Latin divisions. In Section 6, we further clean the rank Latin divisions in order to show that Item iii \( \Rightarrow \) Item vi and Item vii \( \Rightarrow \) Item vii, transposed to the language of ordered graphs. Finally in Section 7, we show that Item viii \( \Rightarrow \) Item vii, transposed to the language of ordered graphs. We also refine the lower bound on the growth of ordered graph classes with unbounded twin-width, to completely settle Balogh et al.'s conjecture [3].

See Figure 2 for a visual outline.

3. Approximating the matrix twin-width is FPT

In this section we show the equivalence between Item i and Item ix. As a by-product, we obtain an \( f(\text{OPT}) \)-approximation algorithm for the twin-width of matrices, or ordered graphs. We first show that a large rich division implies large twin-width. This direction is crucial for the algorithm but not for the main circuit of implications.

Lemma 3.1. If \( M \) has a \( 2k(k+1) \)-rich division \( D \), then \( \text{tww}(M) > k \).

Proof. We prove the contrapositive. Let \( M \) be a matrix of twin-width at most \( k \). In particular, \( M \) admits a \( (k,k) \)-sequence \( P_1, \ldots, P_{n+m-1} \). Let \( D \) be any division of \( M \). We want to show that \( D \) is not \( 2k(k+1) \)-rich.

Let \( t \) be the smallest index such that either a part \( R_i \) of \( P_i^R \) intersects three parts of \( D^R \), or a part \( C_j \) of \( P_i^C \) intersects three parts of \( D^C \). Without loss of generality we can assume that \( C_j \in P_i^C \) intersects three parts \( C_{a,j}, C_{b,j}, C_{c,j} \) of \( D^C \), with \( a < b < c \) where the parts \( C_1, \ldots, C_d \) of the division \( D \) are ordered from left to right. Since \( P_i^C \) is a \( k \)-overlapping partition, the subset \( S \), consisting of the parts of \( P_i^C \) intersecting \( C_{b,j} \), has size at most \( k+1 \). Indeed, \( S \) contains \( C_j \) plus at most \( k \) parts which \( C_j \) is in conflict with.

Here a part \( R_{i,j} \) of \( D^R \) is said red if there exist a part \( R_i \) of \( P_i^R \) intersecting \( R_{i,j} \) and a part \( C_z \) in \( S \) such that the submatrix \( R_i \cap C_z \) is not constant (see Figure 3). We then say that \( C_z \) is a witness of \( R_{i,j} \) being red. Let \( N \subseteq R \) be the subset of rows not in a red part of \( D^R \). Note that for every part \( C_z \in S \), the submatrix \( N \cap C_z \) consists of the same column vector repeated \( |C_z| \) times. Therefore \( N \cap C_{b,j} \) has at most \( k+1 \) distinct column vectors.

Besides, the number of red parts witnessed by \( C_z \in S \) is at most \( 2k \). This is because the number of non-constant submatrices \( R_i \cap C_z \), with \( R_i \in P_i^R \), is at most \( k \) (since \( P_1, \ldots, P_{n+m-1} \) is a \( (k,k) \)-sequence) and because every \( R_i \) intersects at most two parts of \( D^R \) (by definition of \( t \)). Hence the total number of red parts is at most \( 2k|S| \), thus at most \( 2k(k+1) \). Consequently, there is a subset \( X \) of at most \( 2k(k+1) \) parts of \( D^R \), namely the red parts, and a part \( C'_{b,j} \) of \( D^C \) such that \( (R \setminus \cup X) \cap C'_{b,j} = N \cap C'_{b,j} \) consists of at most \( k+1 \) distinct column vectors. Thus \( D \) is not a \( 2k(k+1) \)-rich-division. \( \square \)
Our main algorithmic result is that approximating the twin-width of matrices (or ordered graphs) is FPT. Let us observe that this remains a challenging open problem for (unordered) graphs.

**Theorem 1.2.** Given as input an $n \times m$ matrix $M$ over a finite field $F$, and an integer $k$, there is an $2^{2^{O(k^2 \log k)}} (n + m)^{O(1)}$ time algorithm which returns

- either a $2k(k+1)$-rich division of $M$, certifying that $\text{tww}(M) > k$,
- or an $(|F|^{O(k^4)}, |F|^{O(k^4)})$-sequence, certifying that $\text{tww}(M) = |F|^{O(k^4)}$.

**Proof.** We try to construct a division sequence $D_1, \ldots, D_{n+m-1}$ of $M$ such that every $D_i$ satisfies the following properties $\mathcal{P}^R$ and $\mathcal{P}^C$.

- $\mathcal{P}^R$: For every part $R_a$ of $D_i^R$, there is a set $Y$ of at most $4k(k+1)+1$ parts of $D_i^C$, such that the submatrix $R_a \cap (C \cup Y)$ has at most $4k(k+1)$ distinct row vectors.
- $\mathcal{P}^C$: For every part $C_b$ of $D_i^C$, there is a set $X$ of at most $4k(k+1)+1$ parts of $D_i^R$, such that the submatrix $(R \cup X) \cap C_b$ has at most $4k(k+1)$ distinct column vectors.

The algorithm is greedy: Whenever we can merge two consecutive row parts or two consecutive column parts in $D_i$ so that the above properties are preserved, we do so, and obtain $D_{i+1}$. We first need to show that checking properties $\mathcal{P}^R$ and $\mathcal{P}^C$ are FPT.

**Lemma 3.2.** Deciding if $\mathcal{P}^R$ holds, or similarly if $\mathcal{P}^C$ holds, can be done in time $2^{2^{O(k^2 \log k)}} (n + m)^{O(1)}$.  

**Figure 3.** The division $D$ in black. The column part $C_j \in \mathcal{P}^C_1$, first to intersect three division parts, in orange. Two row parts of $D$ turn red because of the non-constant submatrix $C_z \cap R_i$, with $C_z \in S$ and $R_i \in D^R$. After removal of the at most $2k|S|$ red parts, $|S| \leq k+1$ bounds the number of distinct columns.
Proof. We show the lemma with $\mathcal{R}_R$, since the case of $\mathcal{S}_C$ is symmetric. For every $R_a \in D_i^R$, we denote by $\mathcal{R}_R(R_a)$ the fact that $R_a$ satisfies the condition $\mathcal{R}_R$ starting at “there is a set $Y$.” If one can check $\mathcal{R}_R(R_a)$ in time $T$, one can thus check $\mathcal{R}_R$ and $\mathcal{S}_C$ in time $([D_i^R] + [D_i^C])f(k) \leq (n + m)T$.

To decide $\mathcal{R}_R(R_a)$, we initialize the set $Y$ with all the column parts $C_b \in D_i^C$ such that the zone $R_a \cap C_b$ contains more than $4k(k + 1)$ distinct rows. Indeed these parts have to be in $Y$. At this point, if $R_a \cap (C \cup Y)$ has more than $(4k(k + 1))^{4k(k+1)+2}$ distinct rows, then $\mathcal{R}_R(R_a)$ is false. Indeed, each further removal of a column part divides the number of distinct rows in $R_a$ by at most $4k(k + 1)$. Thus after the at most $4k(k + 1) + 1$ further removals, more than $4k(k + 1)$ would remain.

Let us suppose instead that $R_a \cap (C \cup Y)$ has at most $(4k(k + 1))^{4k(k+1)+2}$ distinct rows. We keep one representative for each distinct row. For every occurring column vector. Now every zone of $R_a \cap C_b$ contains more than $4k(k + 1)$ distinct rows. Thus if that happens, we keep exactly $4k(k + 1) + 2$ copies of $Z$. Thus if that happens, we keep exactly $4k(k + 1) + 2$ copies of $Z$. Now $R_a$ has dimension at most $(4k(k + 1))^{4k(k+1)+2} \times 2^{4k(k+1)}$. Therefore the maximum number of distinct zones is $\exp(\exp(O(k^2 \log k)))$.

If a same zone $Z$ is repeated in $R_a$ more than $4k(k + 1) + 1$ times, at least one occurrence of the zone will not be included in $Y$. In that case, putting copies of $Z$ in $Y$ is pointless: it eventually does not decrease the number of distinct rows. Thus if that happens, we keep exactly $4k(k + 1) + 2$ copies of $Z$. Now $R_a$ has dimension at most $(4k(k + 1) + 2) \cdot \exp(\exp(O(k^2 \log k))) = \exp(\exp(O(k^2 \log k)))$ zones. We can try out all $\exp(\exp(O(k^2 \log k)))^{4k(k+1)+1} \exp(\exp(O(k^2 \log k)))$ possibilities for the set $Y$, and conclude if at least one works.

Two cases can arise.

Case 1. The algorithm terminates on some division $D_i$ and no merge is possible. Let us assume that $D_i^R := \{R_1, \ldots, R_s\}$ and $D_i^C := \{C_1, \ldots, C_t\}$, where the parts are ordered by increasing vector indices.

We consider the division $D$ of $M$ obtained by merging in $D_i$ the pairs $\{R_{2a-1}, R_{2a}\}$ and $\{C_{2b-1}, C_{2b}\}$, for every $1 \leq a \leq \lfloor s/2 \rfloor$ and $1 \leq b \leq \lfloor t/2 \rfloor$. Let $C_j$ be any column part of $D_i^C$. Since the algorithm has stopped, for every set $X$ of at most $2k(k + 1)$ parts of $D_i^R$, the matrix $(R \cup X) \cap C_j$ has at least $4k(k + 1) + 1$ distinct vectors. This is because $2k(k + 1)$ parts of $D_i^R$ corresponds to at most $4k(k + 1)$ parts of $D_i^R$. The same applies to the row parts, so we deduce that $D$ is $2k(k + 1)$-rich. Therefore, by Lemma 3.1, $M$ has twin-width greater than $k$.

Case 2. The algorithm terminates with a full sequence $D_1, \ldots, D_{n+m-1}$. Given a division $D_i$ with $D_i^R := \{R_1, \ldots, R_s\}$ and $D_i^C := \{C_1, \ldots, C_t\}$, we now define a partition $P_i$ that refines $D_i$ and has small error value. To do so, we fix a, say, column part $C_j$ and show how to partition it further in $P_i$.

By assumption on $D_i$, there exists a subset $X$ of at most $r := 4k(k + 1) + 1$ parts of $D_i^R$ such that $(R \cup X) \cap C_j$ has less than $r$ distinct column vectors. We now denote by $F$ the set of parts $R_a$ of $D_i^R$ such that the zone $R_a \cap C_j$ has at least $r$ distinct rows and $r$ distinct columns. Such a zone is said full. Observe that $F \subseteq X$. Moreover, for every $R_a$ in $X \setminus F$, the total number of
distinct column vectors in \( R_a \cap C_j \) is at most \( \max(r, \alpha r^{-1}) = \alpha r^{-1} \), where \( \alpha \geq 2 \) is the size of \( \mathbb{F} \). Indeed, if the number of distinct columns in \( R_a \cap C_j \) is at least \( r \), then the number of distinct rows is at most \( r - 1 \).

In particular, the total number of distinct column vectors in \( (R \setminus \cup F) \cap C_j \) is at most \( w := r(\alpha r^{-1})^2 \); a multiplicative factor of \( \alpha r^{-1} \) for each of the at most \( r \) zones \( R_a \in X \setminus F \), and a multiplicative factor of \( r \) for \( (R \setminus \cup X) \cap C_j \). We partition the columns of \( C_j \) accordingly to their subvector in \( (R \setminus \cup F) \cap C_j \) (by grouping columns with equal subvectors together). The partition \( P_i \) is obtained by refining, as described for \( C_j \), all column parts and all row parts of \( D_i \).

By construction, \( P_i \) is a refinement of \( P_{i+1} \) since every full zone of \( D_i \) remains full in \( D_{i+1} \). Hence if two columns belong to the same part of \( P_i \), they continue belonging to the same part of \( P_{i+1} \). Besides, \( P_i \) is a \( w \)-overlapping partition of \( M \), and its error value is at most \( r \cdot w \) since non-constant zones can only occur in full zones (at most \( r \) per part of \( D_i \)), which are further partitioned at most \( w \) times in \( P_i \). To finally get a contraction sequence, we greedily merge parts to fill the intermediate partitions between \( P_i \) and \( P_{i+1} \). Note that all intermediate refinements of \( P_{i+1} \) are \( w \)-overlapping partitions. Moreover the error value of a column part does not exceed \( r \cdot w \). Finally the error value of a row part can increase during the intermediate steps by at most \( 2w \). All in all, we get a \( (w, (r + 2) \cdot w) \)-sequence. This implies that \( M \) has twin-width at most \( (r + 2) \cdot w = \alpha^{O(k^4)} \).

The running time of the overall algorithm follows from Lemma 3.2.

The approximation ratio, of \( 2^{O(OPT^4)} \), can be analyzed more carefully by observing that bounded twin-width implies bounded VC dimension. Then the threshold \( \alpha r^{-1} \) can be replaced by \( r^d \), where \( d \) upperbounds the VC dimension.

As a direct corollary of our algorithm, if the matrix \( M \) does not admit any large rich division, the only possible outcome is a contraction sequence. Considering the size of the field \( \mathbb{F} \) as an absolute constant, we thus obtain the following.

**Theorem 3.3.** If \( M \) has no \( q \)-rich division, then \( \text{tww}(M) = 2^{O(q^2)} \).

This is the direction which is important for the circuit of implications. The algorithm of Theorem 1.2 further implies that Theorem 3.3 is effective.

4. **Large rich divisions imply large rank divisions**

We remind the reader that a rank-\( k \) division is a \( k \)-division for which every zone has rank at least \( k \). A \((k + 1)\)-rank division is a \( k \)-rich division since the deletion of \( k \) zones in a column of the division leaves a zone with rank at least \( k \), hence with at least \( k \) distinct row vectors. The goal of this section is to provide a weak converse of this statement. We recall that \( \text{mt} \) is the Marcus-Tardos bound of Theorem 2.3. For simplicity, we show the following theorem in the case \( \mathbb{F} = \mathbb{F}_2 \), but the proof readily extends to any finite field by setting \( K \) to \( |\mathbb{F}|^{O(k^4 \text{mt}(k|\mathbb{F}|^k)}) \).

**Theorem 4.1.** Let \( K \) be \( 2^{k \cdot \text{mt}(k|\mathbb{F}|^k)} \). Every 0,1-matrix \( M \) with a \( K \)-rich division \( D \) has a rank-\( k \) division.
Proof. Without loss of generality, we can assume that $\mathcal{D}^C$ has size at least the size of $\mathcal{D}^R$. We color red every zone of $\mathcal{D}$ which has rank at least $k$. We now color blue a zone $R_i \cap C_j$ of $\mathcal{D}$ if it contains a row vector $r$ (of length $|C_j|$) which does not appear in any non-red zone $R_i' \cap C_j$ with $i' < i$. We call $r$ a blue witness of $R_i \cap C_j$.

Let us now denote by $U_j$ the subset of $\mathcal{D}^R$ such that every zone $R_i \cap C_j$ with $R_i \in U_j$ is uncolored, i.e., neither red nor blue. Since the division $\mathcal{D}$ is $K$-rich, if the number of colored (i.e., red or blue) zones $R_i \cap C_j$ is less than $K$, the matrix $(\cup U_j) \cap C_j$ has at least $K$ distinct column vectors. So $(\cup U_j) \cap C_j$ has at least $2^k \text{mt}(k2^k) = \log_2 K$ distinct row vectors. By design, every row vector appearing in some uncolored zone $R_i \cap C_j$ must appear in some blue zone $R_i' \cap C_j$ with $i' < i$. Therefore at least $2^k \text{mt}(k2^k)$ distinct row vectors must appear in some blue zones within column part $C_j$. Since a blue zone contains less than $2^k$ distinct row vectors (its rank being less than $K$), there are, in that case, at least $2^k \text{mt}(k2^k)/2^k = \text{mt}(k2^k)$ blue zones within $C_j$. Therefore in any case, the number of colored zones $R_i \cap C_j$ is at least $\text{mt}(k2^k)$ per $C_j$.

Thus, by Theorem 2.3, we can find $D'$ a $k2^k \times k2^k$ division of $M$, coarsening $D$, with at least one colored zone of $D$ in each cell of $D'$. Now we consider $D''$ the $k \times k$ subdivision of $M$, coarsening $D'$, where each supercell of $D''$ corresponds a $2^k \times 2^k$ square block of cells of $D'$ (see Figure 4). Our goal is to show that every supercell $Z$ of $D''$ has rank at least $k$. This is clearly the case if $Z$ contains a red zone of $D$. If this does not hold, each of the $2^k \times 2^k$ cells of $D'$ within the supercell $Z$ contains at least one blue zone of $D$. Let $Z_{i,j}$ be the cell in the $i$-th row block and $j$-th column block of hypercell $Z$, for every $i,j \in [2^k]$. Consider the diagonal cells $Z_{i,i}$ ($i \in [2^k]$) of $D'$ within the supercell $Z$. In each of them, there is at least one blue zone witnessed by a row vector, say, $\tilde{r}_i$. Let $r_i$ be the prolongation of $\tilde{r}_i$ up until the two vertical limits of $Z$. We claim that every $r_i$ (with $i \in [2^k]$) is distinct. Indeed by definition of a blue witness, if $i < j$, $\tilde{r}_j$ is different from all the row vectors below it, in particular from $r_i$ restricted to these columns. So $Z$ has $2^k$ distinct vectors, and thus has rank at least $k$. 

\hfill \Box

5. Rank Latin divisions

In this section, we show a Ramsey-like result which establishes that every (hereditary) matrix class with unbounded grid rank can encode all the $n$-permutations with some of its $2n \times 2n$ matrices. In particular and in light of the previous sections, this proves the small conjecture for ordered graphs.

We recall that a rank-$k$ $d$-division of a matrix $M$ is a $d$-by-$d$ division of $M$ whose every zone has rank at least $k$, and rank-$k$ $d$-division is a short-hand for rank-$k$ $k$-division. Then a matrix class $\mathcal{M}$ has bounded grid rank if there is an integer $k$ such that no matrix of $\mathcal{M}$ admits a rank-$k$ $k$-division.

Let $I_k$ be the $k \times k$ identity matrix, and $1_k$, $0_k$, $U_k$, and $L_k$ be the $k \times k$ 0,1-matrices that are all 1, all 0, upper triangular, and lower triangular,
Figure 4. In black (purple, and yellow), the rich division $D$. In purple (and yellow), the Marcus-Tardos division $D'$ with at least one colored zone of $D$ per cell. In yellow, the rank-$k$ division $D''$. Each supercell of $D''$ has large rank, either because it contains a red zone (light red) or because it has a diagonal of cells of $D'$ with a blue zone (light blue).

respectively. Let $A^M$ be the vertical mirror of matrix $A$, that is, its reflection about a vertical line separating the matrix in two equal parts. The following Ramsey-like result states that every 0, 1-matrix with huge rank (or equivalently a huge number of distinct row or column vectors) admits a regular matrix with large rank.

**Theorem 5.1.** There is a function $T : \mathbb{N}^+ \rightarrow \mathbb{N}^+$ such that for every natural $k$, every matrix with rank at least $T(k)$ contains as a submatrix one of the following $k \times k$ matrices: $I_k$, $I_k - I_k$, $U_k$, $L_k$, $I_k^M$, $(I_k - I_k)^M$, $U_k^M$, $L_k^M$.

The previous theorem is a folklore result. For instance, it can be readily derived from Gravier et al. [19] or from [9, Corollary 2.4.] combined with the Erdős-Szekeres theorem.

---

8i.e., column $\lceil n/2 \rceil$ if $A$ has $n$ columns and $n$ is odd, and a vertical line between column $n/2$ and $n/2 + 1$ if $n$ is even
Let \( N_k \) be the set of the eight matrices of Theorem 5.1. The first four matrices are said diagonal, and the last four (those defined by vertical mirror) are said anti-diagonal. By Theorem 5.1, if a matrix class \( \mathcal{M} \) has unbounded grid rank, then one can find in \( \mathcal{M} \) arbitrarily large divisions with a matrix of \( N_k \) as submatrix in each row of the division, for arbitrarily large \( k \). We want to acquire more control on the horizontal-vertical interactions between these submatrices of \( N_k \). We will prove that in large rank divisions, one can find so-called rank Latin divisions.

An embedded submatrix \( M' \) of a matrix \( M \) is the matrix \( M' \) together with the implicit information on the position of \( M' \) in \( M \). In particular, we will denote by rows(\( M' \)), respectively cols(\( M' \)) the rows of \( M \), respectively columns of \( M \), intersecting precisely at \( M' \). When we use rows(\( \cdot \)) or cols(\( \cdot \)), the argument is implicitly cast in an embedded submatrix of the ambient matrix \( M \). For instance, rows(\( M \)) denotes the set of rows of \( M \) (seen as a submatrix of itself).

A contiguous (embedded) submatrix is defined by a zone, that is, a set of consecutive rows and a set of consecutive columns. The \((i,j)\)-cell of a \( d \)-division \( D \), for any \( i,j \in [d] \), is the zone formed by the \( i \)-th row block and the \( j \)-th column block of \( D \). A canonical name for that zone is \( D_{i,j} \).

A rank-\( k \) Latin \( d \)-division of a matrix \( M \) is a \( d \)-division \( D \) of \( M \) such that for every \( i,j \in [d] \) there is a contiguous embedded submatrix \( M_{i,j} \in N_k \) in the \((i,j)\)-cell of \( D \) satisfying:

- \( \{\text{rows}(M_{i,j})\}_{i,j} \) partitions rows(\( M \)), and \( \{\text{cols}(M_{i,j})\}_{i,j} \) cols(\( M \)).
- \( \text{rows}(M_{i,j}) \cap \text{cols}(M_{i',j'}) \) equals \( \mathbf{1}_k \) or \( \mathbf{0}_k \), whenever \( (i,j) \neq (i',j') \).

Note that since the submatrices \( M_{i,j} \) are supposed contiguous, the partition is necessarily a 0-overlapping partition, hence a division. A rank-\( k \) pre-Latin \( d \)-division is the same, except that the second item need not be satisfied.

We can now state our technical lemma.

**Lemma 5.2.** For every positive integer \( k \), there is an integer \( K \) such that every 0,1-matrix \( M \) with a rank-\( K \) division has a submatrix with a rank-\( k \) Latin division.

**Proof.** We start by showing the following claim, a first step in the global cleaning process of Lemma 5.2. We recall that \( T(\cdot) \) is the function of Theorem 5.1.

**Claim 5.3.** Let \( M \) be a 0,1-matrix with a rank-\( T(\kappa) \) \( d^2 \)-division \( D \). There is a \( \kappa d^2 \times \kappa d^2 \) submatrix \( \hat{M} \) of \( M \) with a rank-\( \kappa \) \( d \)-division \( \hat{D} \), coarsening \( D \), such that the \((i,j)\)-cell of \( \hat{D} \) contains \( M_{i,j} \in N_\kappa \) as a contiguous submatrix, \( \{\text{rows}(M_{i,j})\}_{i,j} \) partitions rows(\( M \)), and \( \{\text{cols}(M_{i,j})\}_{i,j} \) cols(\( M \)).

**Proof of the claim.** Let \( D^R \) be \( (R_1,\ldots,R_{d^2}) \) and, \( D^C \) be \( (C_1,\ldots,C_{d^2}) \). Let \( D' \) be the coarsening of \( D \) defined by \( D'^R := (\bigcup_{i \in [d]} R_i,\bigcup_{i \in [d+1,2d]} R_i,\ldots,\bigcup_{i \in [(d-1)d+1,d^2]} R_i) \) and \( D'^C := (\bigcup_{i \in [d]} C_j,\ldots,\bigcup_{i \in [(d-1)d+1,d^2]} C_j) \). By Theorem 5.1, each cell of \( D \) contains a submatrix in \( N_\kappa \). Thus there are \( d^2 \) such submatrices in each cell of \( D' \). For every \( i,j \in [d] \), we keep in \( \hat{M} \) the \( \kappa \) rows and \( \kappa \) columns of a single submatrix of \( N_\kappa \) in the \((i,j)\)-cell of \( D' \), and more precisely, one \( M_{i,j} \) in the \((j+(i-1)d,i+(j-1)d)\)-cell of \( D \). In other words, we keep in the \((i,j)\)-cell of \( D' \), a submatrix of \( N_\kappa \) in the \((j,i)\)-cell of
The submatrices $M_{i,j}$ are contiguous in $\tilde{M}$. The set \{rows($M_{i,j}$)$\}_{i,j \in [d]}$ partitions rows($\tilde{M}$) since $j + (i - 1)d$ describes $[d]^2$ when $i \times j$ describes $[d] \times [d]$. Similarly \{cols($M_{i,j}$)$\}_{i,j \in [d]}$ partitions cols($\tilde{M}$).

We denote by $b(k,k)$ the minimum integer $b$ such that every 2-edge coloring of $K_{b,b}$ contains a monochromatic $K_{k,k}$. We set $b^{(1)}(k,k) := b(k,k)$, and for every integer $s \geq 2$, we denote by $b^{(s)}(k,k)$, the minimum integer $b$ such that every 2-edge coloring of $K_{b,b}$ contains a monochromatic $K_{k^2,q^2}$ with $q = b^{(s-1)}(k,k)$. We set $\kappa := b^{(k^2-k^2)}(k,k)$ and $K := \max(T(\kappa), k^2) = T(\kappa)$, so that applying Claim 5.3 on a rank-$K$ division (hence in particular a rank-$T(\kappa) k^2$-division) gives a rank-$\kappa$ pre-Latin $k$-division, with the $k^2$ submatrices of $N_\kappa$ denoted by $M_{i,j}$ for $i,j \in [k]$.

At this point the zones rows($M_{i,j}$) \cap cols($M_{i',j'}$), with $(i,j) \neq (i',j')$, are arbitrary. We now gradually extract a subset of $k$ rows and the $k$ corresponding columns (i.e., the columns crossing at the diagonal if $M_{i,j}$ is diagonal, or at the anti-diagonal if $M_{i,j}$ is anti-diagonal) within each $M_{i,j}$, to turn the rank pre-Latin division into a rank Latin division. To keep our notation simple, we still denote by $M_{i,j}$ the initial submatrix $M_{i,j}$ after one or several extractions.

For every (ordered) pair $(M_{i,j}, M_{i',j'})$ with $(i,j) \neq (i',j')$, we perform the following extraction (in any order of these $\binom{k^2}{2}$ pairs). Let $s$ be such that all

9 Or for readers familiar with the game ultimate tic-tac-toe, at positions of moves forcing the next move in the symmetric cell about the diagonal.
Theorem 6.1

There exists a map $R_t(\cdot) : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that for every $k \geq 1$, $t \geq 1$ the complete graph $K_{R_t(k)}$ with edges colored with $t$ distinct colors contains a monochromatic clique on $k$ vertices, i.e., a clique whose edges all have the same color.

In what follows, for every $p \geq 0$ we denote with $R_t^{(p)}(\cdot)$ the map $R_t(\cdot)$ iterated $p$ times. The core of our proof relies on the following Ramsey-like lemma.

Lemma 6.2. Let $K_{N^2}$ be the complete graph with vertex set $[N]^2$ and $c : E(K_{N^2}) \to [4]$ be a 4-coloring its edges. For every $k \geq 1$, we let $n := R_{16}(k)$ and $q := \binom{n}{2}$. Then if $N \geq R_{16}^{(q+1)}(k)$, there are two subsets $R \in \binom{[N]}{k}$ and $C \in \binom{[N]}{k}$ such that for every $i_1 < i'_1 \in R$, $i_2 < i'_2 \in R$, $j_1 < j'_1 \in C$, $j_2 < j'_2 \in C$:

$$c((i_1,j_1)(i'_1,j'_1)) = c((i_2,j_2)(i'_2,j'_2)), \quad \text{and} \quad c((i_1,j'_1)(i'_1,j_1)) = c((i_2,j'_2)(i'_2,j_2)).$$

Proof. For every pair of rows $i < i' \in [N]$, we define the 16-coloring over the pairs of columns $c_{i,i'} : \binom{[N]}{2} \to [4]^2$ by

$$c_{i,i'}(\{j,j'\}) := (c((i,j)(i',j')), c((i,j')(i',j))).$$

for every $j, j' \in [N]$.

We first let $R' := [n]$ and gradually extract $C' \in \binom{[N]}{n}$ such that for every $i < i' \in R'$, we have $c_{i,i'}(\{j_1,j'_1\}) = c_{i,i'}(\{j_2,j'_2\})$. We denote by $C''$ the set of currently available columns from which we do the next extraction. Initially we set $C'' := [N]$. For every pair $\{i, i'\} \in \binom{R'}{2}$, with $i < i'$, we shrink $C''$ so that $\{\{j,j'\} \mid j \neq j' \in C''\}$ becomes monochromatic with respect to $c_{i,i'}$. More precisely, we iteratively apply Ramsey’s theorem $q$ times. At the start
Lemma 6.4. \( \text{rank Latin division.} \)

Let \( k \geq 1 \) be an integer. Let \( M \) be a matrix with a rank-\( k \) Latin \( N \)-division with \( N := R_{16}^{(q+1)}(k) \), \( q := \binom{n}{2} \), and \( n := R_{16}(k) \). Then there exists \( \eta : \{-1,1\} \times \{-1,1\} \to \{0,1\} \) such that the submatrix closure of \( M \) contains the set \( \{F_{\eta}(\sigma) \mid \sigma \in \mathcal{S}_k\} \).
Proof. Let \((R, C)\) be the rank-\(k\) Latin \(N\)-division, with \(R := \{R_1, \ldots, R_N\}\) and \(C := \{C_1, \ldots, C_N\}\), so that every row of \(R_i\) (resp. column of \(C_i\)) is smaller than every row of \(R_j\) (resp. column of \(C_j\)) whenever \(i < j\). Let \(M_{i,j}\) be the \(\textit{chosen}\) contiguous submatrix of \(N_k\) in \(R_i \cap C_j\) for every \(i, j \in [N]\). We recall that, by definition of a rank Latin division, \(\{\text{rows}(M_{i,j})\}_{i,j \in [N]}\) partitions rows \((M)\) (resp. \(\{\text{cols}(M_{i,j})\}_{i,j \in [N]}\) partitions cols \((M)\)) into intervals.

We now consider the complete graph \(K_{N^2}\) on vertex set \([N]^2\), and color its edges with the function \(c : E(K_{N^2}) \to \{0, 1\}^2\) defined as follows. For every \((i, j) \neq (i', j') \in [N]^2\) (and say, \(i < i'\)), let \(a \in \{0, 1\}\) be the constant entries in rows \(M_{i,j} \cap \text{cols}(M_{i',j'})\), and \(b \in \{0, 1\}\), the constant entries in rows \(M_{i',j'} \cap \text{cols}(M_{i,j})\). Then we define \(c((i, j)(i', j')) := (a, b)\).

We use Lemma 6.2 to find two sets \(R, C \in \binom{[N]}{k}\) such that:

\[
|\{(c((i, j)(i', j')), c((i, j')(i', j'))) \mid i < i' \in R, j < j' \in C\}| = 1.
\]

Let \(\eta : \{-1, 1\} \times \{-1, 1\} \to \{0, 1\}\) be such that \((\eta(-1, -1), \eta(1, 1), \eta(-1, 1), \eta(1, -1)) \in \{0, 1\}^4\) is the unique element of this set. (Note that Lemma 6.2 disregards the edges of \(E(K_{N^2})\) that are between vertices with a common coordinate.) In terms of the rank Latin division, it means that for every \(i < i' \in R\) and \(j < j' \in C\),

- \(\text{cols}(M_{i,j}) \cap \text{rows}(M_{i',j'})\) has constant value \(\eta(-1, -1)\),
- \(\text{rows}(M_{i,j}) \cap \text{cols}(M_{i',j'})\) has constant value \(\eta(1, 1)\),
- \(\text{cols}(M_{i',j'}) \cap \text{rows}(M_{i,j})\) has constant value \(\eta(-1, 1)\), and
- \(\text{rows}(M_{i',j'}) \cap \text{cols}(M_{i,j})\) has constant value \(\eta(1, -1)\).

**Figure 7.** How zones are determined by \(\eta, \text{ot}(i, i')\), and \(\text{ot}(j, j')\).

In other words, \(\text{rows}(M_{i,j}) \cap \text{cols}(M_{i',j'})\) is entirely determined by \(\eta, \text{ot}(i, i')\), and \(\text{ot}(j, j')\) (see Figure 7).

Let \(\sigma \in \Gamma_k\). We now show how to find \(F_{\eta}(\sigma) = (f_{i,j})_{1 \leq i, j \leq k}\) as a submatrix of \(M\). For every \(i \in [k]\), we choose a row \(r_i \in \text{rows}(M_{i,\sigma(i)})\) and a column \(c_{\sigma(i)} \in \text{cols}(M_{i,\sigma(i)})\) such that the entry of \(M\) at the intersection of \(r_i\) and \(c_{\sigma(i)}\) has value \(f_{i,\sigma(i)}\). This is possible since the submatrices \(M_{i,j}\) are in \(N_k\) and have disjoint row and column supports. We consider the \(k \times k\) submatrix \(M'\) of \(M\) with rows \(\{r_i \mid i \in [k]\}\) and columns \(\{c_i \mid i \in [k]\}\).

By design \(M' = F_{\eta}(\sigma)\) holds. Let us write \(M' := (m_{i,j})_{1 \leq i, j \leq k}\) and show for example that if \(\text{ot}(\sigma^{-1}(j), i) = -1\) and \(\text{ot}(j, \sigma(i)) = 1\) for some \(i, j \in [k]\), then we have \(m_{i,j} = \eta(-1, 1) = f_{i,j}\). The other cases are obtained in a similar way. Let \(i' := \sigma^{-1}(j) > i\) and \(j' := \sigma(i) > j\). In \(M'\), \(m_{i,j}\) is obtained by taking the entry of \(M\) associated to the row \(r_i\) of the matrix
$M_{i\sigma(i)} = M_{j\sigma(j)}$ and the column $c_j$ of $M_{\sigma^{-1}(j),j} = M'_{\sigma(j)}$. The entry $m_{i,j}$ lied in $M$ in the zone rows $(M_{i',j'}) \cap \text{cols}(M'_{\sigma(j)})$ with constant value $\eta(-1,1)$. 

We now check that $\sigma \in \mathfrak{S}_k \mapsto F_\eta(\sigma)$ is indeed injective.

**Lemma 6.5.** For every $k \geq 1$ and $\eta : \{-1,1\} \times \{-1,1\} \rightarrow \{0,1\}$:

$$|\{F_\eta(\sigma) \mid \sigma \in \mathfrak{S}_k\}| = k!$$

**Proof.** We let $k \geq 1$ and $\eta : \{-1,1\} \times \{-1,1\} \rightarrow \{0,1\}$. The inequality $|\{F_\eta(\sigma) \mid \sigma \in \mathfrak{S}_k\}| \leq k!$ simply holds. We thus focus on the converse inequality.

When we read out the first row (bottom one) of $F_\eta(\sigma) = (f_{i,j})_{1 \leq i,j \leq k}$ by increasing column indices (left to right), we get a possibly empty list of values $\eta(-1,1)$, one occurrence of $1 - \eta(1,1)$ at position $(1,\sigma(1))$, and a possibly empty list of values $\eta(1,1)$. The last index $j$ such that $f_{1,j} \neq f_{1,j+1}$, or $j = k$ if no such index exists, thus corresponds to $\sigma(1)$. We remove the first row and the $j$-th column and iterate the process on the rest of the matrix. 

We obtain that classes with subfactorial growth have bounded grid rank by piecing Lemmas 5.2, 6.4 and 6.5 together.

**Theorem 6.6.** Every matrix class $\mathcal{M}$ satisfying $|\mathcal{M}_k| < k!$, for some integer $k$, has bounded grid rank.

**Proof.** We show the contrapositive. Let $\mathcal{M}$ be a class of matrices with unbounded grid rank. We fix

$$k \geq 1, \ n := R_{16}(k), \ N := R_{16}^{(16)^{1+1}}(k).$$

Now we let $K := K(N)$ be the integer of Lemma 5.2 sufficient to get a rank-$N$ Latin division. As $\mathcal{M}$ has unbounded grid rank, it contains a matrix $M$ with grid rank at least $K$. By Lemma 5.2, a submatrix $M \in \mathcal{M}$ of $M$ admits a rank-$N$ Latin division, from which we can extract a rank-$k$ Latin $N$-division (since $k \leq N$). By Lemma 6.4 applied to $M$, there exists $\eta$ such that $\{F_\eta(\sigma) \mid \sigma \in \mathfrak{S}_k\} \subseteq \mathcal{M}_k$. By Lemma 6.5 this implies that $|\mathcal{M}_k| \geq k!$. 

We just showed that for every matrix class of unbounded grid rank, for every integer $k$, there is an $\eta(k) : \{-1,1\} \times \{-1,1\} \rightarrow \{0,1\}$ such that $\{F_\eta(k)(\sigma) \mid \sigma \in \mathfrak{S}_k\} \subseteq \mathcal{M}_k \subseteq \mathcal{M}$. As there are only 16 possible functions $\eta$, the sequence $\eta(1), \eta(2), \ldots$ contains at least one function $\eta$ infinitely often. Besides for every $k' < k$, $\{F_\eta(\sigma) \mid \sigma \in \mathfrak{S}_{k'}\}$ is included in the submatrix closure of $\{F_\eta(\sigma) \mid \sigma \in \mathfrak{S}_k\}$. Thus we showed the following more precise result.

**Corollary 6.7.** Let $\mathcal{M}$ be a matrix class with unbounded grid rank. Then there exists $\eta : \{-1,1\} \times \{-1,1\} \rightarrow \{0,1\}$ such that:

$$\mathcal{F}_\eta \subseteq \mathcal{M}.$$
7. Matchings in classes of ordered graphs with unbounded twin-width

We now move to the world of hereditary classes of ordered graphs. In this language, we will refine the lower bound on the slices of unbounded twin-width classes, in order to match the conjecture of Balogh, Bollobás, and Morris [3]. We will also establish that bounded twin-width, NIP, monadically NIP, and tractable (provided that FPT \( \not= AW[*] \)) are all equivalent.

7.1. NIP classes of ordered graphs have bounded twin-width. The following lemma shows how to find encodings of matchings in classes of ordered graphs with unbounded twin-width from the encodings of permutation matrices described in section 6.2.

A crossing function is a mapping \( \eta : \{-1, 1\} \times \{-1, 1\} \cup \{(0, 0)\} \to \{0, 1\} \) with \( \eta(1, 1) \neq \eta(0, 0) \). Let \( \eta \) be a crossing function, let \( n \) be an integer, and let \( \sigma \in \mathcal{S}_n \) be a permutation. We say that an ordered graph \( G \) is an \((\eta, \sigma)\)-matching if \( G \) has vertices \( u_1 < \cdots < u_n < v_1 < \cdots < v_n \) with \( u_iv_j \in E(G) \) if and only if \( \eta(\text{o}(\sigma^{-1}(j), i), \text{o}(j, \sigma(i))) = 1 \). The vertices \( u_1, \ldots, u_n \) and \( v_1, \ldots, v_n \) are respectively the left and the right vertices of \( G \).

Let \( \lambda, \rho : \{-1, 1\} \to \{0, 1\} \) be two mappings. We define \( M_{\eta, \lambda, \rho} \) as the hereditary closure of the class of all \((\eta, \sigma)\)-matchings \( G \) with left vertices \( u_1 < \cdots < u_n \) and right vertices \( v_1 < \cdots < v_n \), such that for every \( 1 \leq i < j \leq n \) we have
\[
\begin{align*}
\text{u_iu_j} \in E(G) & \iff \lambda(\text{o}(\sigma(i), \sigma(j))) = 1, \\
\text{v_iv_j} \in E(G) & \iff \rho(\text{o}(\sigma^{-1}(i), \sigma^{-1}(j))) = 1.
\end{align*}
\]

For fixed \( \lambda, \eta, \rho \), Figure 8 illustrates the rules one have to follow to encode a matching accordingly.

We further define \( M \) as the class of all ordered matchings \( H \) with vertex set \( u_1 < \cdots < u_n < v_1 < \cdots < v_n \), where the matching is between the \( u_i \)'s and the \( v_j \)'s. Note that \( M = M_{\eta, \lambda, \rho} \) for \( \eta \) defined by \( \eta(0, 0) = 1 \) and \( \eta(x, y) = 0 \) if \( (x, y) \neq (0, 0) \), and for \( \lambda \) and \( \rho \) defined by \( \lambda(x) = \rho(x) = 0 \).

**Lemma 7.1.** Let \( \mathcal{C} \) be a hereditary class of ordered graphs with unbounded twin-width. Then there exists a crossing function \( \eta \), such that for every integer \( n \) and every permutation \( \sigma \in \mathcal{S}_n \), the class \( \mathcal{C} \) contains an \((\eta, \sigma)\)-matching.

**Proof.** Let \( M \) be the submatrix closure of the set of adjacency matrices of graphs in \( \mathcal{C} \), along their respective orders. \( M \) has unbounded twin-width (see last paragraph of Section 2.1), and hence unbounded grid rank. By Corollary 6.7, there exists some function \( \eta : \{-1, 1\} \times \{-1, 1\} \to \{0, 1\} \) such that \( \mathcal{F}_\eta \subseteq M \). We may extend the domain of \( \eta \) to \( \{-1, 1\} \times \{-1, 1\} \cup \{(0, 0)\} \) such that it has the desired property.

Let \( \sigma \in \mathcal{S}_n \) be a permutation. Consider its associated matching permutation \( \tilde{\sigma} \in \mathcal{S}_{2n} \) defined by
\[
\tilde{\sigma}(i) := \begin{cases} 
\sigma(i) + n & \text{if } i \leq n \\
\sigma^{-1}(i - n) & \text{if } n + 1 \leq i \leq 2n.
\end{cases}
\]

In other words \( M_{\tilde{\sigma}} \) consists of the two blocs \( M_\sigma \) and \( M_{\sigma^{-1}} \) on its anti-diagonal. We have \( \mathcal{F}_\eta(\tilde{\sigma}) \in M \), so there exists a graph \( H \in \mathcal{C} \) such that
Figure 8. In red, the edges $i\sigma(i)$ of the matching associated to $\sigma \in \mathfrak{S}_n$. On the top drawing, they are crossing, whereas on the bottom one, they are non-crossing. In orange the other edges/non-edges encoded by functions $\lambda, \eta, \rho$. An edge exists in the ordered graph if and only if its label equals 1.

$F_\eta(\tilde{\sigma})$ is a submatrix of its adjacency matrix. Denote by $U_1, U_2$ the (disjoint) ordered sets of vertices corresponding to the rows indexed respectively by $\{1, \ldots, n\}$ and $\{n+1, \ldots, 2n\}$, such that $\max(U_1) < \min(U_2)$. Take similarly $V_1, V_2$ associated to the columns indices. If $\max(U_1) < \min(V_2)$ we let $A = U_1$ and $B = V_2$; otherwise, $\min(U_2) > \max(U_1) \geq \min(V_2) > \max(V_1)$ and we let $A = V_1$ and $B = U_2$. Then, if $u_1 < \cdots < u_n$ are the elements of $A$ and $v_1 < \cdots < v_n$ are the elements of $B$, we have $u_n < v_1$ and $u_iv_j \in E(H)$ if and only if $\eta(\ot(\sigma^{-1}(j), i), \ot(j, \sigma(i))) = 1$. Hence we can let $G = H[A \cup B]$. \hfill \Box

Let $n$ be a positive integer, and let $\sigma \in \mathfrak{S}_n$ be a permutation. A coating permutation of $\sigma$ is a permutation $\varpi \in \mathfrak{S}_{m+n}$ such that $m \geq 2$ and

- $1 = \varpi(1) < \cdots < \varpi(m) = n + m$,
- the pattern of $\varpi$ induced by $[m+1, m+n]$ is $\sigma$, i.e., for every $1 \leq i < j \leq n$ we have $\varpi(i+m) < \varpi(j+m)$ if and only if $\sigma(i) < \sigma(j)$.

The $m$ first vertices are the left coating vertices and their image by $\varpi$ are the right markers.

Lemma 7.2. Let $\eta$ be a crossing function, $\sigma \in \mathfrak{S}_n$, a permutation, $\varpi \in \mathfrak{S}_{n+m}$, a coating permutation of $\sigma$, and $G$, an $(\eta, \varpi)$-matching.

Then the sets of left coating vertices, left vertices, right markers, right vertices, and the matching involution between left coating vertices and right markers are all first-order definable.
Proof. Without loss of generality we assume \( \eta(0,0) = 1 \), for otherwise we can consider \( 1 - \eta \) and the complement of \( G \). In particular, we have \( \eta(1,1) = 0 \). Let \( u_1 < \cdots < u_{n+m} \) (resp. \( v_1 < \cdots < v_{n+m} \)) be the left (resp. right) vertices of \( G \). Let \( 1 \leq i \leq m \). By assumption, if \( 1 < i' < i \) then \( \varpi(i') < \varpi(i) \). Thus (contrapositive, with \( j = \varpi(i') \)) if \( j > \varpi(i) \) then \( \eta^{-1}(j) > i \). As \( \eta(1,1) = 0 \), we deduce that no vertex \( v_j \) with \( j > \varpi(i) \) is adjacent to \( u_i \). As \( \eta(0,0) = 1 \), the vertices \( u_i \) and \( v_{\varpi(i)} \) are adjacent. Hence \( v_{\varpi(i)} \) is definable as the maximum vertex adjacent to \( u_i \). Thus we deduce that (for \( 1 \leq i \leq m \)):

- the vertex \( u_m \) is the minimum vertex adjacent to \( v_{n+m} = \max(V(G)) \) (as \( \varpi(n) = n + m \));
- the left vertices are the vertices that are less or equal to \( u_m \);
- the vertex \( v_{\varpi(i)} \) matched to a left vertex \( u_i \) is the maximum vertex adjacent to \( u_i \);
- a vertex \( v_j \) is a right marker if and only if it is matched to a left vertex, which is then the minimum vertex adjacent to \( v_j \);
- a vertex is a left vertex if it is smaller than \( v_1 \), and a right vertex, otherwise.

\[ \square \]

Lemma 7.3. Let \( \eta \) be a crossing function with \( \eta(0,0) = \eta(1,1) = 1 \). There exists a simple interpretation \( I \) with the following property:

If \( \sigma \in S_n \) is a permutation, \( \varpi \in S_{2n+1} \) is the coating permutation of \( \sigma \) defined by

\[
\varpi(i) := \begin{cases} 
2(i-1) + 1 & \text{if } i \leq n + 1 \\
2\sigma(i - (n + 1)) & \text{if } i > n + 1, 
\end{cases}
\]

and \( G \) is an \( (\eta, \varpi) \)-matching, then \( I(G) \) is the ordered matching defined by \( \sigma \).

Proof. The set of left non-coating vertices and the set of right non-marker vertices are definable according to Lemma 7.2. For a left non-coating vertex \( u_{n+1+i} \), the matching vertex \( v_{2\sigma(i)} \) is the only right non-marker vertex such that the (right marker) vertex just before is non-adjacent to \( u_{n+1+i} \) and the (right marker) vertex just after is adjacent to \( u_{n+1+i} \).

\[ \square \]

Lemma 7.4. Let \( \eta \) be a crossing function with \( \eta(0,0) = \eta(-1,1) = 1 \). There exists a simple interpretation \( I \) with the following property:

If \( \sigma \in S_n \) is a permutation, \( \varpi \in S_{2n+1} \) is the coating permutation of \( \sigma \) defined by

\[
\varpi(i) := \begin{cases} 
2(i-1) + 1 & \text{if } i \leq n + 1 \\
2\sigma(i - (n + 1)) & \text{if } i > n + 1, 
\end{cases}
\]

and \( G \) is an \( (\eta, \varpi^{-1}) \)-matching, then \( I(G) \) is the ordered matching defined by \( \sigma \).

Proof. By interpretation we reverse the ordering of \( G \). This way we get the ordered graph \( G^* \), which is an \( (\eta^*, \varpi) \)-matching, where \( \eta^*(x,y) := \eta(y,x) \). We then apply the interpretation defined in Lemma 7.3.

\[ \square \]

Lemma 7.5. Let \( \eta \) be the crossing function with \( \eta(0,0) = \eta(-1,1) = 1 \), and \( \eta(x,y) = 0 \), otherwise.

There exists a simple interpretation \( I \) with the following property:
If $\sigma \in \mathcal{S}_n$ is a permutation, $\varpi \in \mathcal{S}_{n+2}$ is the only coating permutation of $\sigma$ (with $m = 2$), and $G$ is an $(\eta, \varpi)$-matching, then $I(G)$ is the ordered matching defined by $\sigma$.

**Proof.** By Lemma 7.4, the non-coating left vertices and right non-marker vertices are definable. Let $u$ be a left non-coating vertex and let $v$ be a right non-marker vertex. If $v$ is to the left of the vertex $v'$ matched with $v$ by $\sigma$ then $u$ and $v$ are not adjacent as $\eta(1, 1) = \eta(-1, 1) = 0$. Thus $v'$ is the minimum right non-marker vertex adjacent to $u$. □

**Lemma 7.6.** Let $\eta$ be a crossing function and let $\mathcal{C}$ be a class of ordered graphs containing an $(\eta, \sigma)$-matching for every $\sigma \in \mathcal{S}_n$. Then there exists a simple interpretation $I$ from $\mathcal{C}$ onto $\mathcal{M}$. Moreover, every $n$-edge matching is the interpretation of an ordered graph in $\mathcal{C}$ with at most $4n + 2$ vertices.

**Proof.** This is a direct consequence of the preceding lemmas. □

We deduce:

**Theorem 7.7.** There exists an interpretation $I$, such that for every hereditary class $\mathcal{C}$ of ordered graphs with unbounded twin-width every graph is an $I$-interpretation of a graph in $\mathcal{C}$.

**Proof.** As the class $\mathcal{C}$ is hereditary, there exists a crossing function $\eta$ such that for every permutation $\sigma$ the class $\mathcal{C}$ contains an $(\eta, \sigma)$-matching. Thus we can apply Lemma 7.6 to obtain, by interpretation, a superclass of $\mathcal{M}$.

Before describing the interpretation of graphs in ordered matchings, we show how the ordered matching $M_G$ corresponding to an ordered graph $G$ is constructed.

Let $G$ be an ordered graph with vertices $v_1 < \cdots < v_n$ and edges $e_1, \ldots, e_m$. For $i \in [n]$ and $1 \leq j \leq d(v_i)$ we define $e_{i,j}$ as the index of the $j$th edge incident to $v_i$. The left vertices of $M_G$ will be (in order) $v_1, \ldots, v_n, x, e_1^-, e_1', e_1'^+ , \ldots, e_m^-, e_m', e_m'^+$. The right vertices of $M_G$ will be (in order) $x', e_n, \ldots, e_1, v_1, \ldots, v_n, \epsilon_1, \ldots, \epsilon_j, y, v_j, \epsilon_{j+1}, \ldots, \epsilon_{d(v_j)}$. The matching $M_G$ matches $v_i$ and $e_i'$, $x$ and $x'$, $y$ and $y'$, $e_i$ and $e_i'$, and finally $e_{i,j}$ either with $e_{i,j}^-$ or $e_{i,j}'^+$, depending on whether $v_i$ is the smallest or biggest incidence of $e_{i,j}$ (see Figure 9).

We now prove that there is a simple interpretation $I$, which reconstructs $G$ from $M_G$. First note that $x'$ is definable as the minimum vertex adjacent to a smaller vertex, and $y'$ is definable as the maximum vertex adjacent to a bigger vertex. Also, $x$ is definable from $x'$ and $y$ is definable from $y'$. Now we can define $v_1, \ldots, v_n$ to be the vertices smaller than $x$, ordered with the order of $M_G$. Two vertices $v_i < v_j < x$ are adjacent in the interpretation if there exists an element $e_k > y$ adjacent to a vertex $e_k'$ preceded in the order by an element $e_k^-$ and followed in the order by an element $e_k'^+$ with the following properties: $e_k^-$ is adjacent to a vertex $z^-$ strictly between the neighbor $v_i'$ of $v_i$ and the neighbor of the successor of $v_i$ in the order and, similarly, $e_k'^+$ is adjacent to a vertex $z^+$ strictly between the neighbor $v_j'$ of $v_j$ and the neighbor of the successor of $v_j$ in the order. □

**Corollary 7.8.** Every class $\mathcal{C}$ of ordered graphs with unbounded twin-width is independent.
Theorem 7.9. There exists an interpretation $I$, such that for every (hereditary) class $\mathcal{M}$ of $0,1$-matrices with unbounded twin-width every graph is an $I$-interpretation of a $0,1$-matrix in $\mathcal{M}$.

Proof. Assume $\mathcal{M}$ has unbounded twin-width. Then there exists a crossing function $\eta$ such that $\mathcal{F}_\eta \subseteq \mathcal{M}$.

Let $\mathcal{C} = \mathcal{M}_{\eta,\lambda,\rho}$ where $\lambda$ and $\rho$ are constant functions equal to 0. It follows from Theorem 7.7 that there is an interpretation $I$ such that every graph is an $I$-interpretation of some graph in $\mathcal{C}$.

Let $P$ be the interpretation from $0,1$-matrices to ordered graphs defining $E(x, y) = M(x, y)$. It is clear that $\mathcal{C} = P(\mathcal{F}_\eta)$. Thus every graph is an $I \circ P$-interpretation of a $0,1$-matrix in $\mathcal{M}$. □

Corollary 7.10. Every class $\mathcal{M}$ of $0,1$-matrices with unbounded twin-width is independent.

7.2. Speed jump for classes of ordered graphs. As is, Lemma 7.1 is not powerful enough to obtain the precise value of the speed jump between classes of ordered graphs with bounded and unbounded twin-width, as we have no information about edges in each part of the partition. The following lemma fixes this issue.

Lemma 7.11. Let $\mathcal{C}$ be a hereditary class of ordered graphs. Assume that for every $n \geq 1$ and every induced matching $M$ on $n$ edges, there exists an ordered graph $G \in \mathcal{C}$ and a bipartition $A, B$ of $V(G)$ such that $\max A < \min B$, $|A| = |B| = n$, and $G[A, B]$ is isomorphic to $M$.

Then there is such a graph $G$ further satisfying that adjacencies within $A$ and $B$ are determined by whether the incident edges of $M$ cross or not.
Proof. Let \( n \) be an non-negative integer. We define
\[
    n_0 = R_4(n), \quad n_2 = R_{2^e}(n_0), \quad \text{and} \quad n_1 = R_{2^e}(n_0).
\]
We set \( A := [n_1] \times [n_2] \) and \( B := [\mathbb{F}_2] \times [\mathbb{F}_1] \), where for every integer \( k \), \( \mathbb{F}_k \) denotes a distinct copy of integer \( k \). We consider the perfect matching \((i,j) - (\bar{j},\bar{i})\) between the sets \( A \) and \( B \), and an ordered graph \( G \in \mathcal{C} \) containing it as a semi-induced subgraph.

For \( 1 \leq i < j \leq n_1 \), we color the edge \( ij \) of \( K_{n_1} \) by the isomorphism class of graph \( G[I_i,I_j] \), where \( I_i = \{ i \} \times [n_2] \subseteq A \). Thus we have at most \( 2^{n_2} \) colors. By Ramsey’s theorem, one can therefore find a monochromatic clique \( Z \) of size \( n_0 \) in this colored \( K_{n_1} \). We denote by \( A' \) the set \( \bigcup_{i \in Z} I_i \), and restrict \( B \) to the subset \( B' \) of elements matched with \( A' \). Up to a monotone renaming, we get the perfect matching \((i,j) - (\bar{j},\bar{i})\) between the sets \( A' = [n_0] \times [n_2] \) and \( B' = [\mathbb{F}_2] \times [\mathbb{F}_1] \). We let \( J_i = [\mathbb{F}_2] \times \{ \bar{i} \} \subseteq B' \) and similarly find in \( B' \) a union \( B'' \) of \( n_0 \) sets \( J_i \) such that for every \( J_i, J_j \in B'' \), \( G[J_i,J_j] \) is in the same isomorphism class. Again we let \( A'' \) be the subset of \( A' \) matched to \( B'' \) in \( M \). Without loss of generality we end with a matching \((i,j) - (\bar{j},\bar{i})\) between two copies of \([n_0] \times [n_0]\).

We now define a 4-coloring \( c_{A''} \) of the pairs \( j_1 j_2 \in \binom{[n_0]}{2} \) for \( 1 \leq j_1 < j_2 \leq n_0 \) as follows: for every \( i_1 < i_2 \in [n_0] \) we let
\[
    c_{A''} := (\mathbb{1}_{(i_1,j_1),(i_2,j_2) \in E(G)}, \mathbb{1}_{(i_2,j_1),(i_1,j_2) \in E(G)}).
\]
By our previous extraction in \( A \), this coloring is well defined (it does not depend on the choice of \( i_1 < i_2 \)). By Ramsey’s theorem, there is a subinterval \( I \) of \([n_0]\) inducing a monochromatic clique of size \( n \) in \( K_{n_0} \). We restrict our attention to \( A^{(3)} := I \times [n_0] \subseteq A'' \) and the set \( B^{(3)} \subseteq B'' \) to which \( A^{(3)} \) is matched. We perform the same extraction in \( B^{(3)} \) and obtain \( B^{(4)} \) such that for every \( i_1 < i_2, j_1 < j_2 \), the adjacencies in \( G \) between \((\bar{i_1}, \bar{j_1})\) and \((\bar{i_2}, \bar{j_2})\), and between \((\bar{i_1}, j_2)\) and \((\bar{i_2}, j_1)\) do not depend on the exact values of \( i_1, i_2, j_1, j_2 \). In turn we define \( A^{(4)} \) as the subset of \( A^{(3)} \) matched to \( B^{(4)} \).

Then, given an arbitrary \( n - n \) matching \( M' \), we keep exactly one point in each \( I_i \) of \( A^{(4)} \) and one matched point in each \( J_j \) of \( B^{(4)} \); such that the points realize \( M' \). More precisely if \( \theta \in S_n \) is the permutation associated to \( M' \), we select in \( A^{(4)} \) every vertex \((i, \theta(i))\) and in \( B^{(4)} \) every \((\theta(i), i)\). Now the adjacencies within the left points and within the right points only depend on the fact that the two incident edges of the matching \( M' \) cross. \( \square \)

For the general case we introduce the coding function \( \text{Code}_{\eta} \) associated to a function \( \eta : \{-1,1\} \times \{-1,1\} \cup \{0,0\} \to \{0,1\} \) with \( \eta(1,1) \neq \eta(0,0) \) defined as follows: Let \( G \) be an ordered graph with vertex bipartition \((A,B)\), max \( A < \min B, |A| = |B| = n \), and \( G[A,B] \) be the matching associated to the permutation \( \sigma \in S_n \). We denote by \( v_1 < \cdots < v_n \) the elements of \( A \) and by \( u_1 < \cdots < u_n \) the elements of \( B \). Then \( \text{Code}_{\eta}(G) \) is the ordered graph with vertex set \( A \cup B \), same linear order as \( G \), same adjacencies as \( G \) within \( A \) and within \( B \), and where \( u_i \in A \) is adjacent to \( v_i \in B \) if \( \eta(\sigma^{-1}(j), i), \eta(j, \sigma(i)) \) = 1. It directly follows from Section 6.2 that the coding function \( \text{Code}_{\eta} \) is injective for all admissible \( \eta \). Moreover, the next property is immediate from the definition: Let \( G \) be an ordered graph as
Observe that there are (at most) 256 classes above, let $A$ be any graph on $n$ edges, when considering all possible matchings $M$. By Lemma 7.11, for every matching $M$, we can find an ordered graph $H \in \mathcal{D}$ and two subsets $A'$ and $B'$ of vertices with $|A'| < |B'|$, and $H[A', B']$ isomorphic to $M$, with the property that the adjacencies within $A'$ and $B'$ only depend on the crossing/non-crossing property of the incident edges of $M$. As $A'$ is matched with $B'$, we have $\text{Code}_\eta(H)[A' \cup B'] = \text{Code}_\eta(H[A' \cup B'])$ thus, as $\mathcal{C}$ is hereditary, $\text{Code}_\eta(H[A' \cup B']) \in \mathcal{C}$. As the adjacencies within $A'$ and $B'$ are not changed by $\text{Code}_\eta$, they only depend on the crossing/non-crossing property of the matching hidden by the coding function.

As an immediate consequence we obtain the following:

**Theorem 7.13.** There exist 256 hereditary classes of ordered graphs, namely the $\mathcal{M}_{\eta,\lambda,\rho}$, such that every hereditary class of ordered graphs with unbounded twin-width includes at least one of these classes.

**Proof.** Let $\mathcal{C}$ be a hereditary class of ordered graphs with unbounded twin-width. Lemmas 7.1, 7.11 and 7.12 imply that there exist some crossing function $\eta$ and some mappings $\lambda, \rho : \{-1, 1\} \rightarrow \{0, 1\}$ such that $\mathcal{M}_{\eta,\lambda,\rho} \subseteq \mathcal{C}$. Observe that there are (at most) 256 classes $\mathcal{M}_{\eta,\lambda,\rho}$: one for each triple $\eta, \lambda, \rho$.

We first draw some algorithmic consequence.

**Theorem 7.14.** Assuming $\text{FPT} \neq \text{AW}[\ast]$, FO-model checking is FPT on a hereditary class $\mathcal{C}$ of ordered graphs if and only if $\mathcal{C}$ has bounded twin-width.

**Proof.** Assume $\mathcal{C}$ has unbounded twin-width. We want to show that the existence of a fixed-parameter algorithm $\mathcal{A}$ for first-order model checking on $\mathcal{C}$ would imply the existence of such an algorithm on general (unordered) graphs. If $\text{AW}[\ast] \neq \text{FPT}$ then first-order model checking is not FPT for general graphs, thus it is not FPT on $\mathcal{C}$.

As $\mathcal{C}$ has unbounded twin-width, there is a triple of mappings $\eta^*, \lambda^*, \rho^*$ such that $\mathcal{M}_{\eta^*,\lambda^*,\rho^*} \subseteq \mathcal{C}$. As we do not know $\eta^*, \lambda^*, \rho^*$, we define 256 algorithms $\mathcal{A}_{\eta,\lambda,\rho}$ each of them using $\mathcal{A}$ as a subroutine. One of these algorithms (even if we cannot tell a priori which one) solves the general model checking in fixed-parameter time.

Let $I$ be the interpretation of general graphs in $\mathcal{M}$ and let $J_{\eta,\lambda,\rho}$ be the interpretation of $\mathcal{M}$ in $\mathcal{M}_{\eta,\lambda,\rho}$, for every $\eta, \lambda, \rho$. Let $G$ be any graph on $n$ vertices. We can construct the ordered matching $M \in \mathcal{M}$ such that $I(M) = G$ in time $O(n^2)$. Also in time $O(n^2)$, we can build the 256 ordered graphs
Moreover, let \( H_{\eta,\lambda,\rho} \in \mathcal{M}_{\eta,\lambda,\rho} \) such that \( J_{\eta,\lambda,\rho}(H_{\eta,\lambda,\rho}) = M \), hence \( G = I \circ J_{\eta,\lambda,\rho}(H_{\eta,\lambda,\rho}) \).

Say, we want to check \( G \models \varphi \) for some sentence \( \varphi \) in the language of graphs. There are 256 sentences \( (I \circ J_{\eta,\lambda,\rho})^*(\varphi) \) such that \( G \models \varphi \leftrightarrow H_{\eta,\lambda,\rho} \models (I \circ J_{\eta,\lambda,\rho})^*(\varphi) \), for every \( \lambda, \eta, \rho \). For each of the 256 triples \( \eta, \lambda, \rho \), we define \( A_{\eta,\lambda,\rho} \) as the algorithm which builds \( H_{\eta,\lambda,\rho} \) and then runs \( A \) on the query \( H_{\eta,\lambda,\rho} = (I \circ J_{\eta,\lambda,\rho})^*(\varphi) \).

Among these 256 algorithms is \( A_{\eta^*,\lambda^*,\rho^*} \) which runs in fixed-parameter time, and correctly solves first-order model checking for general graphs. Indeed if \( A \) runs in time \( f(|\varphi|)n^c \) for some computable function \( f \), then \( A_{\eta^*,\lambda^*,\rho^*} \) runs in time \( O(n^2 + g(|\varphi|)n^{2c}) \) for some computable function \( g \).

Now assume that \( \mathcal{G} \) has twin-width at most \( k \). Let \( G \in \mathcal{G} \). Using the fixed-parameter approximation algorithm of Theorem 1.2, we construct a \( 2^O(k) \)-sequence for \( G \) and then apply the FO-model checking algorithm presented in \( [3] \).

### 7.3. Lowerbounding \( |(\mathcal{M}_{\eta,\lambda,\rho})_n| \)

There is still a bit of work to get the exact value of \( \sum_{k=0}^{[n/2]} \binom{n}{2k} k! \) conjectured in \( [3] \) as a lower bound of the growth. We show how to derive this bound in each case of \( \eta, \lambda, \rho \).

We first observe some symmetries to reduce the actual number of cases.

**Lemma 7.15.** For every \( \eta, \lambda, \rho \), \( |(\mathcal{M}_{\eta,\lambda,\rho})_n| = |(\mathcal{M}_{1-\eta,1-\lambda,1-\rho})_n| \).

**Proof.** We simply observe that \( \mathcal{M}_{1-\eta,1-\lambda,1-\rho} \) is the set of (ordered) complements of graphs of \( \mathcal{M}_{\eta,\lambda,\rho} \).

**Lemma 7.16.** For every \( \eta, \lambda, \rho \), \( |(\mathcal{M}_{\eta,\lambda,\rho})_n| = |(\mathcal{M}_{1-\eta,\lambda,\rho})_n| \).

**Proof.** We observe that \( \mathcal{M}_{1-\eta,\lambda,\rho} \) is the set of (ordered) bipartite complements (that is, where one only flips the edges of the bipartition) of graphs of \( \mathcal{M}_{\eta,\lambda,\rho} \).

**Lemma 7.17.** Let \( \eta \) be a crossing function. We define \( \tilde{\eta} \) by \( \tilde{\eta}(x,y) = \eta(y,x) \). Then \( |(\mathcal{M}_{\eta,\lambda,\rho})_n| = |(\mathcal{M}_{\tilde{\eta},\rho,\lambda})_n| \).

**Proof.** The ordered graph corresponding to a permutation \( \sigma \) with the first encoding is obtained from the graph corresponding to \( \sigma^{-1} \) in the second encoding by reversing the linear order.

**Lemma 7.18.** For every integer \( n \geq 0 \), every \( \sigma \in S_n \) and every mappings \( \eta, \lambda, \rho \), \( \mathcal{M}_{\eta,\lambda,\rho} \) contains both the encoding of \( \sigma \) by \( \eta, \lambda, \rho \), and the same graph where all (non-)adjacencies between \( u_i \) and the associated \( v_{\sigma(i)} \) are flipped.

**Proof.** Let \( \sigma^+ \in S_{2n} \) be the permutation defined as follows: For every \( i \in [n] \), \( \sigma^+(2i) = 2\sigma(i) \) and \( \sigma^+(2i - 1) = 2\sigma(i) - 1 \). We encode \( \sigma^+ \) with \( \eta, \lambda, \rho \) and keep only the vertices corresponding to even indices on the left, and to odd vertices on the right. The ordered graph we obtain is the same as the original encoding of \( \sigma \), except that we flipped the adjacencies between the matched vertices. As this new encoding of \( \sigma \) also is in \( \mathcal{M}_{\eta,\lambda,\rho} \), we can conclude.

We observe that the graphs described in the previous lemma constitute a variant of encodings where \( \eta(0,0) \) is allowed to be equal to \( \eta(1,1) \).
Recall that the class $\mathcal{M}$ of ordered matchings is defined as the one $\mathcal{M}_{\eta,\lambda,\rho}$ with $\lambda = \rho = 0$, and $\eta(x,y) = 0$ except $\eta(0,0) = 1$. We denote by $\overline{\mathcal{M}}$ the class of ordered anti-matchings, that is the $\mathcal{M}_{\eta,\lambda,\rho}$ with $\lambda = \rho = 1$, and $\eta(x,y) = 1$ except $\eta(0,0) = 0$. For the classes of ordered matchings and anti-matchings, the bound we want to derive is actually tight.

**Lemma 7.19.** $|\mathcal{M}_n| = \overline{|\mathcal{M}_n|} = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!$.

**Proof.** We only show $|\mathcal{M}_n| = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!$, as Lemma 7.15 implies that $|\mathcal{M}_n| = \overline{|\mathcal{M}_n|}$. The $\binom{n}{2k}$ factor accounts for the number of ways to position the $2k$ matched vertices along $n$ linearly-ordered vertices. The $k!$ counts the number of ways to match, among the $2k$ chosen vertices, the $k$ leftmost ones to the $k$ rightmost ones. Every choice of matched vertices and partial matching gives a distinct ordered graph.

We now deal with $\lambda$ or $\rho$ not being constant.

**Lemma 7.20.** If $\lambda$ or $\rho$ is not constant then $|(\mathcal{M}_{\eta,\lambda,\rho})_n| \geq n! \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!$.

**Proof.** Assume, without loss of generality, that $\lambda$ is not constant. Let $\sigma \in \mathfrak{S}_n$ be any permutation. The permutation $\sigma$ is encoded as an ordered graph $G_\sigma \in \mathcal{M}_{\eta,\lambda,\rho}$ with vertex set $[2n]$ using $\eta$, $\lambda$, and $\rho$. Let $H_\sigma \in \mathcal{M}_{\eta,\lambda,\rho}$ be the restriction of $G_\sigma$ to $[n]$. As $\lambda(1) \neq \lambda(-1)$ we can retrieve all the inversions of $\sigma$ in $[n]$ from the ordered graph $H_\sigma$, thus we can retrieve $\sigma$ as well. It follows that $\sigma \mapsto H_\sigma$ is an injection from $\mathfrak{S}_n$ into $(\mathcal{M}_{\eta,\lambda,\rho})_n$ hence $|(\mathcal{M}_{\eta,\lambda,\rho})_n| \geq n!$.

Now we deal with the remaining cases.

**Lemma 7.21.** For every encoding mappings $\eta, \lambda, \rho$ such that $\lambda$ and $\rho$ are constant, and either $\lambda \neq \rho$ or $\lambda$ takes value $\eta(1,1)$, we have for every integer $n \geq 1$,

$$|\mathcal{M}_{\eta,\lambda,\rho}_n| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!.$$  

**Proof.** We fix $n \geq 1$ and the encoding mappings $\eta, \lambda, \rho$. By Lemma 7.15 we may assume that $\lambda$ is constant with value 1.

For every $k \in [n]$, $\sigma \in \mathfrak{S}_k$, and $X \in \binom{\{a_1, \ldots, a_k\}}{k}$, we partition $X = A \cup B$ into the set $A = \{a_1 < \cdots < a_k\}$ of its $k$ smallest elements and $B = \{b_1 < \cdots < b_k\}$ the set of its $k$ largest elements. We observe that $b_1 \geq k+1$ since $a_1, \ldots, a_k$ are $k$ distinct integers in $[n]$ all smaller than $b_1$. Our goal is to construct a permutation $\sigma_{(A,B)} \in \mathfrak{S}_{n-k}$, encoding that $\sigma$ is applied precisely between $A$ and $B$. We will partition $[n]$ into two intervals: the vertices of index at most $b_1 - 1$ and the vertices of index at least $b_1$. The permutation $\sigma_{(A,B)}$ matches $A$ and $B$ according to $\sigma$, and the rest of the vertices with “ancillary vertices” in a way that helps identifying the position of the “primary vertices” (that is, vertices of $A \uplus B$).

We now detail the construction. For every $i \in [b_1 - 1]$, we have a vertex $u_i$. These $b_1 - 1$ vertices are ordered $u_1 < u_2 < \cdots < u_{b_1-1}$, and form a set denoted by $U$. For every $i \in [b_1, n]$, we have a vertex $v_i$. These $n-b_1+1$ vertices are ordered $v_{b_1} < v_{b_1+1} < \cdots < v_n$, and form a set denoted by $V$. 
We add, for every \( i \in [b_1, n] \setminus B \), a vertex \( u'_i \). These \( n - b_1 + 1 - k \) vertices are ordered by increasing indices, and form a set called \( U' \). Finally we add, for every \( i \in [b_1 - 1] \setminus A \), a vertex \( v'_i \). These \( b_1 - 1 - k \) vertices are ordered by increasing indices, and form a set called \( V' \).

\[
U := U' \uplus U, \quad V := V \uplus V'
\]

with the total orders inherited from the ones on \( U, U', V, V' \) and the relations \( \max(U') < \min(U) \) and \( \max(V) < \min(V') \).

Moreover we order the set \( U \uplus V \) with the relation \( \max(U) < \min(V) \). Note that all the vertices of \( U \) are “to the left” of all the vertices of \( V \) and that both these sets have \( n - k \) elements. The disjoint sets \( U \) and \( V \) may be identified as a bipartition set \([n]\). In turn \( A \) and \( B \) may be identified as \( k \)-subsets of \( U \) and \( V \), respectively. The sets \( U' \) and \( V' \) are extra vertices necessary to match the vertices of \( V \setminus B \) and \( U \setminus A \). Now we define the matching permutation \( \sigma_{(A,B)} \) between \( U \) and \( V \) as follows:

\[
\sigma_{(A,B)}(u) := \begin{cases} 
  v_{b \sigma(j)} & \text{if } u = u_j \text{ with } i = a_j \in A \\
  v'_i & \text{if } u = u_i \text{ with } i \notin A \\
  v_i & \text{if } u = u'_i.
\end{cases}
\]

Intuitively this matching encodes \( \sigma \) between the copies of \( A \) and \( B \) in \( U \) and \( V \), and matches \( U \setminus A \) to \( V' \), and \( U' \) to \( V \setminus B \), in an order-preserving fashion.

Now we show that this encoding is injective, i.e., that for every \( k, k' \geq 1 \), \( \sigma, \sigma' \in \mathfrak{S}_k \), \( A = A \uplus B \in \binom{[n]}{2k} \) and \( X = [n] \setminus \{a_k\} \), \( X' = [n] \setminus \{a_k\} \), if \( M, M' \) denote respectively the encodings under \( \eta, \lambda, \rho \) of \( \sigma_{(A,B)} \) and \( \sigma'_{(A',B')} \), then

\[
M[U \uplus V] \approx M'[U \uplus V] \Rightarrow k = k', \quad \sigma = \sigma', \quad (A, B) = (A', B'),
\]

where \( H \approx H' \) means that (ordered) graph \( H \) is isomorphic to (ordered) graph \( H' \). (Note that, as we presently deal with totally ordered graphs, the isomorphism is imposed by the linear orders and straightforward to find.)

We consider \( M[U \uplus V] \) for an encoding \( M \) of \( \sigma_{(A,B)} \), and show that we can deduce the values of \( k, \sigma, A \) and \( B \) from it. First we show that we can find the maximum \( u_{b_1 - 1} \) of \( U \) by the assumptions made on the mappings \( \eta, \lambda, \rho \). If \( \lambda \) is constant to \( \eta(1, 1) = 1 \), then \( \eta(0, 0) = 0 \) and \( u_{b_1 - 1} \) is the largest vertex \( u \) of \( M[U \uplus V] \) which is adjacent with all the vertices \( w < u \).

If \( \lambda \) and \( \rho \) are constant with different values, then \( \rho = 0 \), and \( u_{b_1 - 1} \) is simply the only vertex of \( M[U \uplus V] \) non-adjacent to its successor but adjacent to its predecessor, except in the very special case where \( \max(A) = b_1 - 1 = a_k \) is matched with \( \min(B) = b_1 \) (i.e. \( \sigma(k) = 1 \)).

We now deal with this special case. If \( \eta(-1, 1) = 0 \), then \( u_{b_1 - 1} \) is the maximum vertex of \( U \uplus V \) forming a clique with all the vertices “to its left.” If \( \eta(1, -1) = 1 \), then \( u_{b_1 - 1} \) is the maximum vertex of \( U \uplus V \) not forming an independent set with the vertex “to its right.” The other cases reduce to these two by Lemma 7.17.

Hence we can identify \( u_{b_1 - 1} \) from the restriction \( M[U \uplus V] \in \mathcal{M}_{\eta, \lambda, \rho} \). If \( b_1 - 1 \notin A \), then there is an edge between \( u_{b_1 - 1} \) and the vertices \( v_i \in V \) whenever \( \eta(1, 1) = 1 \), by construction of \( \sigma_{(A,B)} \). Otherwise if \( b_1 - 1 \in A \), then there is an edge between \( u_{b_1 - 1} \) and its image by \( \sigma \), namely \( v_{b \sigma(k)} \), whenever \( \eta(0, 0) = 1 \) (hence \( \eta(1, 1) = 0 \)). Hence we can determine whether or not \( b_1 - 1 \) is in \( A \). Moreover when \( b_1 - 1 \in A \), since \( u_{b_1 - 1} \) is the maximum
of $U$, the adjacencies between $u_{b_1-1}$ and every vertex $v_j$ with $j < b_{\sigma(k)}$ are all the same, determined by $\eta(1, 1)$, hence we can find $v_{\sigma(k)}$. If we remove only $u_{b_1-1}$ in the first case, or $u_{b_1-1}$ together with $v_{b_{\sigma(k)}}$ in the second case, then we can iteratively determine all the sets $A$ and $B$ and uniquely build the permutation $\sigma$ between them. Hence we proved the injectivity of our encoding.

This implies that there are $\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!$ distinct such ordered graphs $M[U \cup V]$, which all belong to $(\mathcal{M}_{\eta,\lambda,\rho})_n$, hence we get the desired result. \hfill \Box

We finally slightly tune the previous proof to cover the rest of the cases.

**Lemma 7.22.** For every encoding mappings $\eta, \lambda, \rho$ such that $\lambda$ and $\rho$ are constant and equal, if $\eta(x, y) = \lambda(1)$ for some $x, y \in \{-1, 1\}$, then we have for every $n \geq 1$:

$$|(\mathcal{M}_{\eta,\lambda,\rho})_n| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!.$$ 

**Proof.** By Lemma 7.15 we may assume that $\lambda = \rho = 1$. If $\eta(1, 1) = 1$, then we are done by Lemma 7.21. Thus we may safely assume that $\eta(1, 1) = 0$. By Lemma 7.18 we will only consider ordered graphs obtained by removing the possible edges at matched pairs from the encoding of $\eta, \lambda, \rho$.

Now further assume that $\eta(1, -1) = 1$. We repeat the construction of Lemma 7.21 for every $k \geq 0$, $\sigma \in \mathfrak{S}_k$ and every pair $(A, B)$, but this time we “cut” earlier between the “left” and “right” vertices. We now want $a_k$ as the maximum of $U$ (and the minimum of $V$ may not be in $B$). Moreover, this time we place $V'$ to the left of $V$, that is, we let $\max(V') < \min(V)$. Following the previous proof, we get the injectivity this time by “reading the matching from right to left.” Indeed if we consider $v := \max(V)$, then either $v \notin B$ and we detect it as it is adjacent to every other vertex, or $v \in B$ and we detect it as it is non-adjacent to some previous vertex. Moreover, the vertex it is matched to is the maximum vertex not adjacent to $v$. Hence we may proceed as before.

By Lemma 7.17 we are also done when $\eta(-1, 1) = 1$.

Finally we assume that $\eta(-1, -1) = 1$. We do the same construction as in Lemma 7.21 (cut between $b_1-1$ and $b_1$), and this time we place $U'$ to the right of $U$ and $V'$ to the left of $V$, i.e., we impose $\max(U) < \min(U')$ and $\max(V') < \min(V)$. Similar arguments apply again, and we obtain the injectivity by reading the vertices “from left to right.” \hfill \Box

We can now conclude.

**Theorem 7.23.** For every $\eta, \lambda, \rho$ and every $n \geq 1$:

$$|(\mathcal{M}_{\eta,\lambda,\rho})_n| \geq \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} k!.$$ 

**Proof.** By Lemmas 7.20 to 7.22 we are done unless $\lambda$ and $\rho$ are constant and equal, and $\eta$ is constant on $\{-1, 1\} \times \{-1, 1\}$ with the opposite value to $\lambda$ and $\rho$. By Lemma 7.15 we thus can assume that $\lambda = \rho = 0$, and $\eta(x, y) = 1$ for every $x, y \in \{-1, 1\}$. Now we apply the reduction of Lemma 7.16 and obtain the triple of mappings $\eta, \lambda, \rho$ with $\lambda = \rho = 0$, and $\eta(x, y) = 0$ for every
\(x, y \in \{-1, 1\}\) (thus \(\eta(1, 1) = 0\)). This is the class of ordered matchings, so we conclude by Lemma 7.19. □

We leave as an open question to exhibit a Ramsey-minimal family of ordered graph classes with unbounded twin-width.

**Acknowledgments.** We thank Eunjung Kim, Jarik Nešetřil, Sebastian Siebertz, and Rémi Watrigant for fruitful discussions.

**References**


