The treewidth of proofs☆

Moritz Müller

Universität Wien
Kurt Gödel Research Center for Mathematical Logic
Währinger Straße 25,1090 Vienna, Austria
moritz.mueller@univie.ac.at

Stefan Szeider

Technische Universität Wien
Algorithms and Complexity Group
Favoritenstraße 9-11, 1040 Vienna, Austria
stefan@szeider.net

Abstract

So-called *ordered* variants of the classical notions of pathwidth and treewidth are introduced and proposed as proof theoretically meaningful complexity measures for the directed acyclic graphs underlying proofs. Ordered pathwidth is roughly the same as proof space and the ordered treewidth of a proof is meant to serve as a measure of how far it is from being treelike. Length-space lower bounds for k-DNF refutations are generalized to arbitrary infinity axioms and strengthened in that the space measure is relaxed to ordered treewidth.

Keywords: Proof Complexity, Infinity Axioms, Treewidth, Pathwidth, Resolution, Proof Space.

---

1. Introduction

Razborov says that "in most cases the basic question of propositional proof complexity boils down to this. Given a mathematical statement encoded as a propositional tautology ϕ and a class of admissible mathematical proofs formalized as a propositional proof system P, what is the minimal possible complexity of a P-proof of ϕ?" [41, p.415] This is also the perspective of "Bounded Reverse Mathematics" taken in Cook and Nguyen's monograph [13, p.xiv].

1.1. Resolution-based proof systems

A proof system of fundamental interest is Resolution. The most important complexity measures for refutations are the *length*, the *width* and the *space* of a resolution refutation. Space (formula-space or clause-space) has been introduced by Esteban and Torán [18]. Intuitively, a space 100 refutation of a set Γ of clauses is one that can be presented as follows.

> A teacher is in class equipped with a blackboard containing up to 100 clauses. The teacher starts from the empty blackboard and finally arrives at one containing the empty clause. The blackboard can be altered by either writing down a clause from Γ, or by wiping out some clause, or by deriving a new clause from clauses currently written on the blackboard by means of the Resolution rule.

Some interesting restrictions of Resolution are obtained by requiring a particular simple structure of the DAGs (directed acyclic graphs) underlying refutations. Examples are *Input*, *linear* and *treelike* Resolution – we refer to the monograph [26]. Interesting extensions of Resolution include $R(1), R(2), \ldots, R(\log)$ from [28]. The system $R(1)$ is

☆An extended abstract of this work appeared as [35].

just Resolution, and $R(k)$ is a straightforward generalization operating with k-DNFs instead of clauses. The treelike versions of these systems are all simulated by (daglike) Resolution [27], so all treelike and daglike systems $R(k)$ line up in a hierarchy. The hierarchy is strict with respect to length as shown in [17] for the treelike systems and in [46, 44] for the daglike ones. The hierarchy is also strict with respect to space, see again [17] for the treelike, and [7] for the daglike systems.

From a practical perspective the special interest in Resolution derives from its connections to SAT-solvers with length and space of refutations corresponding to time and space of algorithms. We refer to [37] for a recent survey. From the more theoretical perspective of "Bounded Reverse Mathematics", the systems $R(k)$ deserve some special interest because length lower bounds for them imply independence from weak arithmetics based on various forms of $\forall\exists$-induction schemes. For example, super-quasipolynomial length lower bounds on treelike or daglike $R(\log)$ imply independence from relativized bounded arithmetics $T^1_2(\alpha)$ or $T^2_2(\alpha)$ respectively [28]. See [9] for independence derivable from super-polynomial length lower bounds for daglike $R(1), R(2), \ldots$

Concerning the relationship of the complexity measures for (daglike) Resolution, Ben-Sasson and Wigderson [8] famously showed how to derive length lower bounds from width lower bounds. Also space lower bounds follow from width lower bounds [2] (see [19] for a recent alternative proof) but not vice-versa [36]. Ben-Sasson [6] initiated "the research of optimizing two of the measures at once" [6] and proved a *trade-off*, i.e. a negative answer, for length and width in treelike Resolution. Recently, Razborov [42] found an "ultimate" such trade-off. Ben-Sasson and Nordström [7] proved various trade-offs for length and space, for example, they constructed CNFs refutable by (daglike) Resolution in length $O(n)$ as well as in space $O(n/\log n)$, but every refutation in this space has length $2^{n^{\Omega(1)}}$. Beame et al. [5] found a length-space trade-off applying to Resolution refutations of superlinear space.

1.2. Infinity axioms

Many of the abovementioned lower bounds for the different complexity measures are witnessed by quite artificial CNFs. Recalling the introductory quote, CNFs that naturally express certain combinatorial principles deserve some special interest. A large class of such CNFs is obtained from first-order sentences φ letting CNFs $\langle\varphi\rangle_n$ naturally describe models of φ of size n. If φ does not have finite models, then these CNFs are contradictory and we ask for the complexity to refute them. If φ has no model at all, there are polynomial length refutation even in treelike Resolution [43]. If φ has no finite but an infinite model, i.e., φ is an *infinity axiom*, then exponential length lower bounds have been shown for the treelike systems, namely $2^{\Omega(n)}$ for treelike Resolution by Riis [43], $2^{\Omega(n\log k/k)}$ for treelike $R(k)$ by Dantchev and Riis [16], and already earlier $2^{\Omega(\sqrt{n})}$ for treelike $R(\log)$ by Krajíček [29].

But the daglike systems have short refutations of some infinity axioms. Stålmarck [47] gave a polynomial length Resolution refutation of the (CNFs expressing the) *least number principle*, the infinity axiom asserting a pre-order without minimal elements. Dantchev and Riis [16] showed that Resolution needs exponential length to refute any *relativized* infinity axiom. Iterating relativizations of the least number principle yields natural witnesses to the exponential separations of $R(k)$ and $R(k+1)$ [15]. It is not understood which (say, by some model-theoretic criterion) infinity axioms do have short refutations, say, in $R(k)$ for constant k; see [14] for a discussion.

As a second example, Maciel et al. [32] gave quasipolynomial length $R(\log)$-refutations of the *weak pigeonhole principle* with n^2 pigeons and n holes. It is not known whether this can be improved to polynomial. A lower bound $2^{\Omega(n/(\log n)^2)}$ is known [40] for Resolution. We refer to [39, 45] for surveys of the proof complexity of pigeonhole principles.

For Resolution, space lower bounds have been obtained in [18] for the pigeonhole principles and in [1] for the least number principle. [17] generalizes these bounds to $R(k)$.

1.3. Ordered treewidth

Short $R(\log)$-refutations of infinity axioms cannot be treelike, in Razborov's words, they "must necessarily use a high degree of parallelism." [42, Abstract]. It would be desirable to quantify the amount of parallelism used by a proof and consider it as a complexity measure of proofs.

A hint how to do so comes from considering space. Space can be seen as a connectivity measure of the DAG underlying a refutation: Esteban and Torán [18] characterized space as a certain pebbling number of the refutation DAG. Following Beame et al. [5] the space of a linearly written Resolution refutation is the minimal number w such that at any derivation step at most w many already derived clauses are to be used at a later step. These characterizations are superficially reminiscent of characterizations of pathwidth for undirected graphs (see [24] and [25]), the second being akin to the vertex separation number.

Pathwidth and treewidth play an important role in Robertson and Seymour's graph minors project and have evolved as very successful and ubiquitously used complexity measures of graphs. We refer to [10] for a survey. Many graph problems can be efficiently solved by dynamic programing on a tree-decomposition witnessing small treewidth (see e.g. [20, Chapter 11]), and in fact treewidth turned out to be the key parameter to understand the complexity of graph homomorphism problems ([22, 33, 31] is a sample of some seminal results).

With an eye to proof DAGs, we introduce notions of path and tree decompositions of digraphs with associated width notions *ordered pathwidth* and *ordered treewidth*. Starting with [23] a number of width notions for digraphs have been proposed (see [21] for a survey, [4] for a monograph). But, citing Kreutzer and Ordyniak, "all digraph decompositions proposed so far measure in some way the similarity of a graph to being acyclic. In particular, acyclic graphs have small width in all of these measures" [30, p.4689]. In contrast, the *ordered* width notions allow us to distinguish between DAGs. We feel that these notions are handy in that they enjoy some of the basic combinatorics familiar from the undirected setting. The notions are well-motivated from a graph theoretic point of view; for example on DAGs, ordered pathwidth coincides with a straightforward variant of the vertex separation number adapted to DAGs (Proposition 3.17).

More importantly, we show that the ordered width notions have proof theoretic sense: the connection to the vertex separation number readily implies that the pathwidth of a refutation DAG is roughly the same as its space (Proposition 5.1). Resolution refutations of minimal ordered pathwidth 'are' Input Resolution refutations (up to some elementary rewriting, see Theorem 4.1 and its proof). Conceptually, this allows us to think of space as a measure of how far a Resolution refutation is from being an Input Resolution refutation.

We propose ordered treewidth as a measure of parallelism, that is, of how far a refutation is from being treelike. We show that Resolution refutations of minimal ordered treewidth 'are' treelike refutations (see Theorem 4.3). We also give an interpretation of ordered treewidth in terms of space (Theorem 5.4), using the following two player game, that continues the metaphor above.

> A student visits the teacher in her office asking her to explain the proof. The teacher has a blackboard potentially containing up to 10 clauses and writes the empty clause on it. The student asks how to prove it. The teacher produces a length ≤ 10 proof from Γ plus some additional clauses. The student chooses one of these additional clauses and asks how to prove it. The blackboard is cleaned, the teacher answers and so on. The game ends when the teacher comes up with a proof using no additional clauses.

1.4. Lower bounds

Our main result (Theorem 6.1) is a lower bound on length and ordered treewidth for $R(k)$-refutations of infinity axioms in general. More precisely, let k, w, ℓ be functions of n and φ an infinity axiom; then $R(k)$-refutations of $\langle\varphi\rangle_n$ of ordered treewidth w and length ℓ must satisfy

$$k \cdot w \cdot \log \ell \geq n^{\Omega(1)}.$$

This generalizes the lower bounds for treelike $R(\log)$ mentioned above. It makes progress with respect to the known length-space lower bounds in that it applies to infinity axioms in general, and thereby to a large class of formulas having a natural meaning. It relaxes the refutation space measure (i.e., ordered pathwidth) to ordered treewidth, and it gives nontrivial lower bounds for all $R(k)$ simultaneously, and for $R(\log)$. The latter feature overcomes a bottleneck in constructions from [7] which give good lower bounds for $R(k)$ with constant k but become trivial for $R(\log)$.

We state some corollaries concerning issues mentioned in this introduction. First, as a corollary to the proof, we also get lower bounds on space, i.e., ordered pathwidth, for infinity axioms in general. Namely, $R(k)$-refutations of $\langle\varphi\rangle_n$ of ordered pathwidth w and any length must satisfy $k \cdot w \geq n^{\Omega(1)}$ (Corollary 6.3).

Concerning short refutations of infinity axioms we can now make quantitative sense of the statement that they require a high degree of parallelism, even for the rather strong system $R(\log)$. Namely, already subexponential length $2^{n^{o(1)}}$ $R(\log)$-refutations of $\langle\varphi\rangle_n$ require ordered treewidth $n^{\Omega(1)}$ (Corollary 6.6).

This implies a trade-off for length and parallelism that is witnessed by a natural example, namely the least number principle, and where the upper bounds hold for Resolution while the lower bound holds for $R(\log)$ (Corollary 6.8).

1.5. Proof idea

The proof of our main result follows the adversary type argument of [29] against treelike $R(\log)$-refutations of infinity axioms. One uses restrictions that describe finite parts of some infinite model of the infinity axiom. Starting with the empty restriction, first choose a node as in Spira's theorem, namely one that splits the refutation tree into two subtrees of size at most 2/3 of total. Then distinguish two cases, namely whether no extension of the current restriction satisfies the formula at the chosen node or not. In the first case stick with the current restriction and recurse to the subtree rooted at the chosen node. In the second case delete this subtree and recurse with a "small" extension of the current restriction that satisfies the formula at the chosen node.

The invariant maintained is a proof of a formula "forced false" from axioms plus some formulas "forced true." If the proof has length ℓ, this process reaches a constant size proof after $O(\log \ell)$ steps. If ℓ is not too large, it is argued that the final restriction can be further extended to force all remaining axioms true and a contradiction is reached.

The proof of our lower bound proceeds similarly but by recursion on an ordered tree decomposition of the refutation. To make sense of this idea we show that we can always find a tree decomposition whose underlying tree is binary (to find a Spira type split node) and whose size is linear in the length of the refutation (Lemma 3.12). Further care is needed to ensure that the partial tree decompositions during the recursion are decompositions of refutations with similar properties as the invariant described above (Lemma 3.8).

2. Preliminaries

2.1. Digraphs

We consider *directed graphs* (*digraphs*, for short) without self-loops and denote the set of vertices and the set of directed edges of a digraph D by $V(D)$ and $E(D)$, respectively. If $(u, v) \in E(D)$, then u is a *predecessor* of v and v a *successor* of u. An *ancestor* of $v \in V(D)$ is a vertex w such that there is a directed path from w to v in D; we understand that vertices in a path are pairwise distinct, and that there is a directed path from any vertex to itself. By the *length* of a path we mean its number of edges. The *in-degree (out-degree) of* v is the number of its predecessors (successors). The in-degree (out-degree) of D is the maximal in-degree (out-degree) over all vertices. Vertices of in-degree 0 are *sources*, vertices of out-degree 0 are *sinks*. An *(induced) subdigraph* of D is a digraph $D[X]$ induced on a nonempty $X \subseteq V(D)$; if $V(D) \setminus X$ is nonempty, we write $D - X$ for $D[V(D) \setminus X]$. The graph $\underline{D}$ *underlying* a digraph D has the same vertices as D and as edges $E(D) \cup \{(u, v) \mid (v, u) \in E(D)\}$, the symmetric closure of $E(D)$. In general, a *graph* is a digraph D with symmetric $E(D)$. A *DAG* is a directed acyclic graph (i.e., a digraph without directed cycles), and a *tree* is a DAG T with a unique sink r_T called *root* such that for every $v \in V(T)$ there is exactly one directed path from v to r_T. We shall refer to vertices in a tree as *nodes*. The *subtree* T_t *rooted at* $t \in V(T)$ is the subtree of T induced on the set of ancestors of t in T; it has root $r_{T_t} = t$. The *height* of a tree is the maximal length in a *branch* (leaf-to-root path) in T. By the *perfect binary tree* B_h of height h we mean the tree where every node which is not a leaf has exactly two predecessors and all branches have length exactly h.

2.2. Propositional logic

A *literal* is a propositional variable X or its negation $\neg X$; for a literal ℓ we let $\neg\ell$ denote $\neg X$, if $\ell = X$, and X, if $\ell = \neg X$. A *(k-)term* is a set of (at most k) literals. A *(k-)DNF* is a set of (k-)terms. The empty DNF is denoted by 0 and the empty term by 1. A *clause* is a 1-DNF. An *assignment* is a function from the propositional variables into $\{0, 1\}$. A *restriction* ρ is a finite partial assignment. For a restriction or assignment ρ and a term t we let $t \upharpoonright \rho$ be 0 if t contains a literal falsified by ρ (in the usual sense) and otherwise the subterm obtained by deleting all literals satisfied by ρ. For a DNF F we let $F \upharpoonright \rho$ be 1 if $t \upharpoonright \rho = 1$ for some $t \in F$; otherwise $F \upharpoonright \rho$ is the DNF $\{t \upharpoonright \rho \mid t \in F\} \setminus \{0\}$. Note that, if ρ is defined on all variables appearing in F then $F \upharpoonright \rho$ equals the truth value of D under ρ.

Definition 2.1. A *(k-)DNF proof* is a pair $(D, (F_v)_{v \in V(D)})$ where D is a DAG with a unique sink and in which every vertex has at most two predecessors, and F_v is a (k-)DNF for every $v \in V(D)$. The proof is said to be *of* F if $F = F_v$ for v the sink of D, and *from* Γ if $F_v \in \Gamma$ for all sources v of D. It is said to be *treelike* if D is a tree. Proofs of 0 are *refutations*. The *length* of the proof is $|V(D)|$. A *refutation system* is a set of refutations.

Usually one requires refutation systems to satisfy certain further properties like soundness or completeness or being polynomial time decidable (cf. [12]).

Definition 2.2. A DNF proof $(D,(F_v)_{v\in V(D)})$ is *sound* if for every inner vertex $v \in V(D)$ and every assignment ρ we have $F_v \restriction \rho = 1$ whenever $F_u \restriction \rho = 1$ for all predecessors u of v in D. It is *strongly sound* (cf. [46]) if for every inner vertex $v \in V(D)$ and every restriction ρ we have $F_v \restriction \rho = 1$ whenever $F_u \restriction \rho = 1$ for all predecessors u of v in D.

Obviously, strongly sound proofs are sound. The next statement is also obvious.

Lemma 2.3. *If there is a strongly sound proof of F from Γ and ρ is a restriction such that $G \restriction \rho = 1$ for all $G \in \Gamma$, then $F \restriction \rho = 1$.*

We consider the following rules of inference, namely *weakening*, *introduction of conjunction* and *cut*:

$$\frac{F}{F \cup \{t\}} \qquad \frac{F \cup \{t\} \quad F' \cup \{t'\}}{F \cup F' \cup \{t \cup t'\}} \qquad \frac{F \cup \{t\} \quad F' \cup F''}{F \cup F'},$$

where F, F', F'' are DNFs, t, t' are terms and in the cut rule we assume $\emptyset \neq F'' \subseteq \{\{\neg\ell\} \mid \ell \in t\}$. A k-DNF proof $(D,(F_v)_{v\in V(D)})$ is an $R(k)$-proof if for every inner vertex v with predecessors u, w the formula F_v is obtained from F_u and F_w by one of the three rules above. An $R(k)$-proof is an $R(\log)$-*proof* if its length is at least 2^k. An $R(1)$-proof is a *Resolution* proof. The refutation system consisting of all $R(k)$-refutations (resp. $R(\log)$-refutations) is denoted $R(k)$ (resp. $R(\log)$).

Remark 2.4. $R(k)$ is strongly sound. We have completeness in the sense that for every k-DNF F implied by some set Γ of k-DNFs, there is an $R(k)$-proof of F from Γ plus some additional 'axioms' of the form $(X \vee \neg X)$, i.e., $\{\{X\},\{\neg X\}\}$. $R(k)$ is refutation-complete in the sense that no such axioms are needed in case $F = 0$. If one adds a new rule allowing to infer such an axiom from any formula, then the system ceases to be strongly sound.

2.3. *First-order logic and propositional translation*

A *vocabulary* is a finite set τ of relation and function symbols, each with an associated *arity*; function symbols of arity 0 are *constants*. The *arity* of τ is the maximum arity of one of its symbols. τ-*terms* are first-order variables $x, y, z \ldots$ or of the form $ft_1 \cdots t_r$ where the t_i are again τ-terms and $f \in \tau$ is an r-ary function symbol. τ-*atoms* are of the form $t_1{=}t_2$ or $Rt_1 \cdots t_r$ where the t_i are τ-terms and $R \in \tau$ an r-ary relation symbol. τ-*formulas* are built from τ-atoms using $\wedge, \vee, \neg$ and quantification $\exists x, \forall x$. For a tuple of first-order variables $\bar{x}$ we write $\varphi(\bar{x})$ for a τ-formula φ to indicate that the free variables of φ are among the components of $\bar{x}$. A τ-*sentence* is a τ-formula without free variables. A τ-*structure* M consists of a nonempty *universe*, that we also denote by M, and for all r-ary relation and function symbols $R \in \tau$ and $f \in \tau$ *interpretations* $R^M \subseteq M^r$ and $f^M : M^r \to M$; we identify the interpretation of a constant with its unique value. A τ-structure M is a *model of* a τ-sentence φ if φ is true in M.

The *spectrum* of a first-order sentence φ is the set of those naturals $n \geq 1$ such that φ has a model (with universe) of cardinality n. An *infinity axiom* is a satisfiable first-order sentence with empty spectrum, i.e., a sentence without a finite but with an infinite model. Skolemization and elementary formula manipulation allows to compute from every first-order sentence ψ a sentence φ with the same spectrum and of a special form defined below; moreover, φ has an infinite model if and only if ψ does.

Definition 2.5. A first-order formula is *ready for translation* if it is of the form

$$\forall \bar{x} \textstyle\bigwedge_{i\in I} C_i(\bar{x}),$$

where I is a nonempty finite set and the C_is are first-order clauses (disjunctions of atoms and negated atoms) whose atoms have the form $R\bar{y}$ or $f\bar{y}{=}z$ or $y{=}z$ for some relation symbol R, function symbol f and variables $\bar{y}, y, z$.

Example 2.6. The *least number principle* asserts a pre-order without minimal elements. It is formulated using a unary function symbol f and a binary relation symbol $<$:

$$lnp := \forall xyz(\neg x{<}x \wedge (\neg x{<}y \vee \neg y{<}z \vee x{<}z) \wedge (\neg fx{=}y \vee y{<}x)).$$

Example 2.7. The *weak functional pigeonhole principle* asserts that n^2 pigeons fly injectively into n holes. This is formulated using a binary function symbol f:

$$wphp := \forall xx'yy'z((\neg fxx'{=}z \vee \neg fyy'{=}z \vee x{=}y) \wedge (\neg fxx'{=}z \vee \neg fyy'{=}z \vee x'{=}y')).$$

Following Paris and Wilkie (cf. [38], see also [43]) we define for every natural $n > 0$ a set $\langle\varphi\rangle_n$ of clauses that is satisfied exactly by those assignments that describe a model of φ with universe

$$[n] := \{0, 1 \ldots, n-1\}.$$

Let φ be a τ-sentence ready for translation. We use the expressions $R\bar{b}$ and $f\bar{b}{=}c$ as propositional variables where $r \in \mathbb{N}, \bar{b} \in [n]^r, c \in [n]$, R is an r-ary relation symbol in τ and f is an r-ary function symbol in τ. A truth assignment of these variables describes τ-structures with universe $[n]$ provided it satisfies the *functionality clauses* $\{\{f\bar{b}{=}c\} \mid c \in [n]\}$ and $\{\{\neg f\bar{b}{=}c\}, \{\neg f\bar{b}{=}c'\}\}$ for $f \in \tau$ an r-ary function symbol, $\bar{b} \in [n]^r$ and distinct $c, c' \in [n]$.

The set $\langle\varphi\rangle_n$ contains these functionality clauses and further clauses $\langle C_i(\bar{a})\rangle$ for all $i \in I$ and $\bar{a} \in [n]^{|\bar{x}|}$. To define $\langle C_i(\bar{a})\rangle$, substitute $\bar{a}$ for $\bar{x}$ in $C_i(\bar{x})$; this transforms every literal into a propositional literal or into an expression of the form $a{=}a'$ or $\neg a{=}a'$ where a, a' are components of $\bar{a}$; the propositional clause $\langle C_i(\bar{a})\rangle$ is $\{1\}$ if one of these expressions is "true" in the obvious sense; otherwise $\langle C_i(\bar{a})\rangle$ is the clause whose terms are the singletons of the propositional literals (of the form $(\neg)R\bar{b}, (\neg)f\bar{b}{=}c$) obtained by the substitution.

Remark 2.8. Refuting $\langle\varphi\rangle_n$ needs width, in fact so-called Poizat-width, $n^{\Omega(1)}$ in any strongly sound refutation system whatsoever [3]. However, this does not imply lower bounds on size or space using the mentioned results of [8, 2] because $\langle\varphi\rangle_n$ contains clauses of width at least n and at least n^2 many variables. This follows from the fact that every infinity axiom φ that is ready for translation contains at least one function symbol of positive arity (see Remark 6.2).

3. Width notions for DAGs

3.1. Treewidth and pathwidth

Let G be graph. A *tree decomposition* of G is a pair (T, χ) where T is a tree and χ is a function from $V(T)$ into the powerset of $V(G)$ such that:

(a) every vertex of G belongs to $\chi(t)$ for some $t \in V(T)$;
(b) for every edge $(v, w) \in E(G)$ there is $t \in V(T)$ such that $v, w \in \chi(t)$;
(c) for every $v \in V(G)$ the set $\{t \in V(T) \mid v \in \chi(t)\}$ is connected in $\underline{T}$.

Recall, $\underline{T}$ is the graph underlying T. The *width* of a tree decomposition (T, χ) is the maximum $|\chi(t)| - 1$ over all $t \in V(T)$. The *treewidth* $tw(G)$ of G is the minimum width over all its tree decompositions. A *path decomposition* is a tree decomposition (T, χ) where T is a (directed) path. The *pathwidth* $pw(G)$ of a graph G is the minimum width over all its path decompositions.

Let (T, χ) be a tree decomposition of a graph G. We say that a vertex $v \in V(G)$ is *introduced* at $t \in V(T)$ if $v \in \chi(t)$ but $v \notin \chi(t')$ for all predecessors t' of t. Similarly, we say that v is *forgotten* at $t \in V(T)$ if $v \in \chi(t)$ and either $t = r_T$ or $v \notin \chi(t')$ for the successor t' of t. Note that every vertex $v \in V(G)$ is introduced at least one tree node (by condition (a)) and forgotten at exactly one tree node (by condition (c)). In a path decomposition every vertex is introduced at exactly one tree node.

The same definitions apply literally to digraphs, so we can also speak of tree and path decompositions of digraphs. Consequently, the treewidth and pathwidth of a digraph equal the treewidth and pathwidth of the digraph's underlying graph, respectively. Thus the direction of edges is completely irrelevant for the treewidth or pathwidth of a digraph. For some considerations, however, one needs the direction of edges to be reflected in the decomposition and the associated width measure. For example [23] introduces the notion of *directed treewidth*, and it is known that every DAG has directed treewidth 1. We introduce new width measures that can distinguish between DAGs.

3.2. Ordered treewidth and ordered pathwidth

Although we shall be mainly interested in DAGs, we give the definitions and some first observations generally for digraphs.

Definition 3.1. A tree decomposition (T, χ) of a digraph D is *ordered* if the following condition holds:

(d) for every directed edge $(u, v) \in E(D)$ and every $t \in V(T)$ where v is introduced, $u \in \chi(t)$.

As above, we define the *ordered treewidth* $otw(D)$ of D as the minimum width over all ordered tree decompositions of D, and the *ordered pathwidth* $opw(D)$ of D as the minimum width over all ordered path decompositions of D.

We say that a class C of digraphs has *bounded ordered pathwidth* if there is a constant $w \in \mathbb{N}$ such that every digraph in C has ordered pathwidth at most w; we say C has *unbounded ordered pathwidth* if it does not have bounded ordered pathwidth. We use a similar mode of speech for the other width notions.

Remark 3.2. For every digraph D, $otw(D)$ is at least the in-degree of D.

Examples 3.3.

1. The ordered treewidth of a tree (with edges directed towards the root) is its in-degree.
2. A directed path with at least one edge has ordered pathwidth 1.
3. The class of perfect binary trees (with edges directed towards the root) has unbounded ordered pathwidth and bounded ordered treewidth.
4. The class of perfect binary trees with all edges reversed (edges directed away from the root) has unbounded ordered treewidth and bounded treewidth.

Proof of (1)–(3). (1) and (2). A tree (path) T has the ordered (path) tree decomposition (T, χ) where χ maps $t \in V(T)$ to the set containing t and its predecessors. It has minimal width by Remark 3.2.

(3). Recall B_h denotes the the perfect binary tree of height h (see Section 2.1). By (1) $otw(B_h)$ is 2 for $h > 0$ and 0 for $h = 0$. It is well-known that $pw(B_h) \geq \lceil h/2 \rceil$ (see, e.g., [10, Theorem 67]). This implies (3) noting $opw \geq pw$. $\square$

We prove (4) after Lemma 3.11 below.

The following two lemmas show that ordered treewidth or pathwidth is not increased by taking "minors" in a certain sense (more restrictive than the one in [23, Section 5]).

Lemma 3.4. *Let D be a digraph, (T, χ) an ordered tree decomposition of D and $X \subseteq V(D)$ be nonempty. Then (T, χ') is an ordered tree decomposition of $D[X]$ where χ' maps $t \in V(T)$ to $\chi(t) \cap X$.*

We omit the straightforward proof.

Lemma 3.5. *Let D be a DAG and (T, χ) an ordered tree decomposition of D. Assume $v \in V(D)$ has in-degree 1 and predecessor u and let D' be obtained by contracting the edge (u, v), i.e., by deleting v and adding edges from u to the successors of v. Then (T, χ') is an ordered tree decomposition of D', where for $t \in V(T)$*

$$\chi'(t) := \begin{cases} \chi(t) & \text{if } v \notin \chi(t); \\ (\chi(t) \setminus \{v\}) \cup \{u\} & \text{otherwise.} \end{cases}$$

Proof. Evidently (T, χ') satisfies conditions (a) and (b) of a tree decomposition. To verify condition (c), we need to show that the set $U = \{ t \in V(T) \mid u \in \chi'(t) \}$ is connected in $\underline{T}$. By construction, U is the union of the sets $\{ t \in V(T) \mid u \in \chi(t) \}$ and $\{ t \in V(T) \mid v \in \chi(t) \}$ which are both connected in $\underline{T}$ since (T, χ) is a tree decomposition. The sets share a node t with $u, v \in \chi(t)$, hence U is connected in $\underline{T}$.

It remains to verify condition (d). For edges $(w, x) \in E(D')$ with $x \neq u$ the condition clearly holds since (T, χ) is ordered. Hence consider an edge $e = (w, u) \in E(D')$ and let $t \in V(T)$ such that u is introduced in (T, χ') at t. We observe that $u \in \chi(t)$, since otherwise v would be introduced at t in (T, χ) without u being in $\chi(t)$, contradicting that (T, χ) is ordered. Hence u is introduced at t also in (T, χ), and thus $w \in \chi(t)$ and $w \in \chi'(t)$ as required. $\square$

In the previous lemma, the assumption that v has in-degree 1 can not be omitted:

Example 3.6. A star with n vertices and all edges directed towards the center can be obtained from B_h by contracting edges provided h is sufficiently large. Then $otw(B_h) = 2$ while the star has ordered treewidth $n - 1$.

Definition 3.7. A subtree T' of a tree T is *fully in T* if for every node of T' either all or none of its predecessors in T are in $V(T')$.

Lemma 3.8. *Let (T, χ) be an ordered tree decomposition of a digraph D, let T' be a subtree of T and set $\chi' := \chi \upharpoonright V(T')$. Assume that $\bigcup_{t' \in V(T')} \chi(t') \neq \emptyset$ and set $D' := D[\bigcup_{t' \in V(T')} \chi(t')]$. Then*

1. *(T', χ') is an ordered tree decomposition of D';*
2. *if T' is fully in T, then there exists for every edge $(u, v) \in E(D)$ with $u \notin V(D')$ and $v \in V(D')$ a leaf t of T' which is not a leaf of T such that $v \in \chi(t)$.*

Proof. (1). That (T', χ') satisfies conditions (a) and (c) is easy to see. To verify (b), let $(u, v) \in E(D')$ and choose $t_u, t_v \in V(T')$ such that $u \in \chi(t_u)$ and $v \in \chi(t_v)$. By condition (b) for (T, χ) we find $t_{uv} \in V(T)$ such that $u, v \in \chi(t_{uv})$. Choose ℓ minimal such that there is a path $t_1 \cdots t_\ell$ in $\underline{T}$ with $t_1 = t_{uv}$ and $t_\ell \in V(T')$. Then every path in $\underline{T}$ from t_{uv} to some node in $V(T')$ contains t_ℓ. In particular, this holds for all paths in $\underline{T}$ connecting t_u and t_{uv}. Then $u \in \chi(t_\ell)$ since (T, χ) satisfies condition (c). Similarly $v \in \chi(t_\ell)$, and (b) for (T', χ') follows. Thus (T', χ') is a tree decomposition of D'.

We verify condition (d), i.e., that (T', χ') is ordered. Let $(u, v) \in E(D')$ and assume v is introduced at t_1 in (T', χ'). We have to show $u \in \chi(t_1)$. In (T, χ), the vertex v must be introduced at some ancestor t_2 of t_1, that is, at some $t_2 \in V(T_{t_1})$. Since (T, χ) is ordered, $u \in \chi(t_2)$. We already verified (b) for (T', χ'), so there must be a node $t_3 \in V(T')$ with $u, v \in \chi'(t_3)$. If $t_3 = t_1$ we are done, so assume $t_3 \neq t_1$. Then t_3 cannot be an ancestor of t_1 since (T', χ') satisfies (c) and v is introduced at t_1 in (T', χ'). Hence $t_3 \in V(T) \setminus V(T_{t_1})$. Then the path between t_2 and t_3 in $\underline{T}$ contains t_1. By condition (c) for (T, χ) then $u \in \chi(t_1)$.

(2). Assume T' is fully in T and let $(u, v) \in E(D)$ with $u \notin V(D')$ and $v \in V(D')$. Choose $t' \in V(T')$ such that $v \in \chi(t')$. In (T, χ), v is introduced at some ancestor t of t'. Then $u \in \chi(t)$ because (T, χ) is ordered. Since $u \notin V(D')$, we have $t \notin V(T')$. In (T, χ), v is contained in every bag on the directed path from $t \notin V(T')$ to $t' \in V(T')$, and in particular, in the bag of the first node $t'' \in V(T')$ that we reach on this path. Then t'' is not a leaf of T (since t'' has ancestor $t \neq t''$). It also has some predecessor outside $V(T')$, namely its predecessor on the mentioned path. Since T' is fully in T, all predecessors of t'' are outside $V(T')$, i.e. t'' is a leaf of T'. □

Definition 3.9. A tree decomposition (T, χ) is *succinct* if every node forgets some vertex.

Lemma 3.10. *A succinct ordered tree decomposition of a digraph D has at most $|V(D)|$ many nodes.*

Proof. Let (T, χ) be an ordered tree decomposition of D. As already mentioned, every vertex of D is forgotten at exactly one node of T. This defines a function from $V(D)$ into $V(T)$. Succinctness of (T, χ) means that this function is surjective. □

Lemma 3.11. *Every digraph D has a succinct ordered tree decomposition of width otw(D), and a succinct ordered path decomposition of width opw(D).*

Proof. We only prove the first statement. Let (T, χ) be a width $otw(D)$ ordered tree decomposition of D with the smallest number of nodes. We claim (T, χ) is succinct. Assume there is a node $s \in V(T)$ that does not forget some vertex. It suffices to construct a new tree decomposition (T', χ') with $\chi' := \chi \upharpoonright V(T')$ where $V(T') = V(T) \setminus \{s\}$. If $s = r_T$, then $\chi(r_T) = \emptyset$. In this case, r_T has predecessors $t_1, \ldots, t_r$ for some $r > 0$. We define T' by (declaring t_1 to be the new root and) adding edges (t_i, t_1) for $1 < i \leq r$. If $s \neq r_T$, then s has a successor t in T with $\chi(s) \subseteq \chi(t)$. In this case we define T' by adding all edges (t', t) for $(t', s) \in E(T)$. □

Proof of Examples 3.3 (4). Write B_h^{-1} for B_h with all edges reversed. Clearly, $tw(B_h^{-1})$ is 1 for $h > 0$ and 0 for $h = 0$. For $h > 0$ we show that

$$otw(B_h^{-1}) \geq \tfrac{1}{2} \log h. \tag{1}$$

By Lemma 3.11 there exists a succinct ordered tree decomposition (T, χ) of B_h^{-1} of minimal width $w := otw(B_h^{-1})$.

Claim 1. T has at most $2^{w+1} - 1$ many leaves.

Proof of Claim 1. For every leaf t of T let $N_t \subseteq \chi(t)$ be the set of vertices forgotten at t; this set is nonempty by succinctness. Consider a leaf t of T and a vertex $v \in N_t$. Because t is a leaf and the decomposition is ordered, $\chi(t)$ contains all ancestors of v in B_h^{-1}. Since $|\chi(t)| \leq w + 1$, it follows that v has at most $w + 1$ ancestors in B_h^{-1}, hence v is of distance at most w from the root r_{B_h} of B_h. Now, B_h^{-1} has exactly $2^{w+1} - 1$ vertices that are of distance at most w from the root. Since each such vertex can occur in at most one set N_t for a leaf t, the claim follows. ⊣

Claim 2. $pw(B_h^{-1}) < (2^{w+1} - 1)(w + 1)$.

Proof of Claim 2. Let P be a longest branch in T, and let $t_1, \ldots, t_m = r_T$ be its nodes in order. For $i \in [m]$ let $\chi'(t_i)$ be the union of all sets $\chi(s)$ where $s \in V(T)$ is a node in T of distance exactly $m - i$ from the root $t_m = r_T$. By Claim 1 there are at most $2^{w+1} - 1$ such nodes s. Then (P, χ') is a path decomposition (in fact, even an ordered one) of B_h^{-1} and has width at most $(2^{w+1} - 1)(w + 1) - 1$. ⊣

As mentioned in Examples 3.3 (3), $pw(B_h^{-1}) = pw(B_h) \geq \lceil h/2 \rceil$, so we have $h < 2(2^{w+1} - 1)(w + 1) < 2^{2w+2}$ by Claim 2. This implies (1). □

Lemma 3.12. *For every digraph D there exists an ordered tree decomposition (T, χ) of width $otw(D)$ where T has in-degree at most 2 and $|V(T)| < 2|V(D)|$.*

Proof. By Lemmas 3.11 and 3.10, any DAG D has a an ordered tree decomposition (T, χ) of width $otw(D)$ and $|V(T)| \leq |V(D)|$. As long as there are *bad* nodes of in-degree at least 3 repeat the following. Choose a bad node t and two of its predecessors t_0, t_1; delete edges $(t_0, t), (t_1, t)$, add a new node s and add edges $(s, t), (t_0, s), (t_1, s)$; give s the bag $\chi(t) \cap (\chi(t_0) \cup \chi(t_1))$.

The result is again a tree decomposition of D. To see it is ordered note that the new node s does not introduce any vertices, and the set of vertices introduced at t does not change.

The procedure terminates because each repetition decreases by 1 the sum of in-degrees of bad nodes. The procedure adds at most one new node per edge of T, so the final tree decomposition has at most $|V(D)| + (|V(T)| - 1) < 2|V(D)|$ many nodes. □

Proposition 3.13. *Let $w, \ell \geq 1$ and (T, χ) a width w ordered tree decomposition of a digraph D such that T has height ℓ. Then $opw(D) < (w + 1) \cdot (\ell + 1)$.*

Proof. By adding if necessary some nodes with empty bags we can assume that all branches of T have the same length, say, ℓ. If we order $V(T)$ in an arbitrary way, then every branch naturally corresponds to a tuple from $[d]^\ell$ where d is the in-degree of T. Then branches are ordered via the lexicographical order on $[d]^\ell$. Use the path of branches according to this order as the path underlying a path decomposition. The bag at the ith path node is the union of the $\ell + 1$ many bags $\chi(t)$ for t ranging over the ith branch in T. It is straightforward to verify that this defines an ordered path decomposition of D. The size of bags is bounded by $(w + 1) \cdot (\ell + 1)$. □

3.3. Vertex separation numbers

A *linear layout* (or *linear arrangement*) of a graph G with n vertices is a bijection $\phi : V(G) \to [n]$. For every $i \in [n]$ we define four sets of vertices:

$$\begin{aligned}
L_G(i, \phi) &:= \{u \in V(G) \mid \phi(u) \leq i\}, \\
R_G(i, \phi) &:= \{u \in V(G) \mid \phi(u) > i\}, \\
L_G^*(i, \phi) &:= \{u \in L_G(i, \phi) \mid \exists v \in R_G(i, \phi) : (u, v) \in E(G)\}, \\
R_G^*(i, \phi) &:= \{v \in R_G(i, \phi) \mid \exists u \in L_G(i, \phi) : (u, v) \in E(G)\}.
\end{aligned}$$

The *in-degree* and the *out-degree* of ϕ is defined as $\max_{i \in [n-1]} |R_G^*(i, \phi)|$ and respectively $\max_{i \in [n-1]} |L_G^*(i, \phi)|$. The *vertex separation number $vsn(G)$* of G is defined as the smallest out-degree over all linear layouts of G (which equals the smallest in-degree over all linear layouts of G).

Proposition 3.14 ([24])**.** *$pw(G) = vsn(G)$ for every graph G.*

Note that the definition of in-degree and out-degree of a linear layout makes sense for digraphs. Recalling that a digraph is a DAG if and only if there exist linear layouts such that all (directed) edges run from left to right, it is natural to consider the following variant of the vertex separation number (for DAGs):

Definition 3.15. A linear layout ϕ of a DAG D is *ordered* if for every $(u, v) \in E(D)$ we have $\phi(u) < \phi(v)$. The *ordered vertex separation number $ovsn(D)$* of a DAG D is the smallest out-degree over all ordered linear layouts of D.

We note that this definition is not symmetric in the sense that, in general, if we replace "smallest out-degree" by "smallest in-degree" we get a different number.

Example 3.16. Let D be a star with n vertices and all edges directed towards the center v. Then every ordered linear layout ϕ satisfies $\phi(v) = n$, has in-degree 1 and out-degree $n - 1$.

We prove an analogue of Proposition 3.14.

Proposition 3.17. *$opw(D) = ovsn(D)$ for every DAG D.*

Proof. Let D be a DAG with n vertices. First we show that $opw(D) \geq ovsn(D)$. Let (T, χ) be an ordered path decomposition of D of width w. Let $t_1, \dots, t_m$ be the vertices of T given in the ordering as we visit them when traversing T from the leaf to the root. Recall that every vertex of D is introduced at exactly one node in T. Let $\psi : V(D) \to [m]$ be the function such that a vertex v of D is introduced at node $t_{\psi(v)}$. We define the inverse ϕ^{-1} of a linear layout ϕ of D recursively as follows. To determine $\phi^{-1}(j)$, let r be minimal such that $N_r := \chi(t_r) \setminus \{\phi^{-1}(i) \mid i < j\} \neq \emptyset$; choose a source v (say, the smallest according to some fixed order of $V(G)$) of the DAG $D[N_r]$ induced on N_r, and set $\phi^{-1}(j) := v$.

To see that the linear layout ϕ is ordered, consider a directed edge $(u, v) \in E(D)$. Let $\phi(v) = j$ and $\psi(v) = r$. We have $u \in \chi(t_{\psi(v)})$ since (T, χ) is ordered, so $\psi(u) \leq \psi(v)$. If $\psi(u) < \psi(v)$, then clearly $\phi(u) < \phi(v)$. If $\psi(u) = \psi(v)$, then in the above process we assign $\phi(u)$ before we assign $\phi(v)$, hence $\phi(u) < \phi(v)$ as well.

To see that the out-degree of ϕ is at most w, consider $L^*_D(i, \phi)$ for some $i \in [n - 1]$. Let $v = \phi^{-1}(i + 1)$ and consider a vertex $u \in L^*_D(i, \phi)$. By definition, u has a successor $u' \in R_D(\phi, i)$, and clearly $\phi(u) \leq \phi(v) \leq \phi(u')$. This implies $\psi(u) \leq \psi(v) \leq \psi(u')$. Since ϕ is ordered and $(u, u') \in E(D)$ it follows that $u \in \chi(t_{\psi(u')})$, and by definition, $u \in \chi(t_{\psi(u)})$. By condition (c) for (T, χ), it follows that $u \in \chi(t_j)$ for all $\psi(u) \leq j \leq \psi(u')$, and in particular $u \in \chi(t_{\psi(v)})$. Thus $L^*_D(i, \phi) \subseteq \chi(t_{\psi(v)})$. Moreover, we have $v \in \chi(t_{\psi(v)}) \setminus L^*_D(i, \phi)$. Thus $L^*_D(i, \phi) \subseteq \chi(t_{\psi(v)}) \setminus \{v\}$, and hence $|L^*_D(i, \phi)| \leq |\chi(t_{\psi(v)}) \setminus \{v\}| \leq w$.

Next we show that $opw(D) \leq ovsn(D)$. Let $\phi : V(D) \to [n]$ be an ordered layout of D with out-degree w. We define a path decomposition (T, χ) letting T be the directed path $([n], \{(i, i + 1) \mid i \in [n - 1]\})$ and setting $\chi(i) := L^*_D(i - 1, \phi) \cup \{\phi^{-1}(i)\}$ for $i \in [n]$; here, we understand $L^*_D(i - 1, \phi) = \emptyset$ for $i = 0$. It is easy to verify that (T, χ) is a path decomposition of D and each bag has size at most $w + 1$. To see it is ordered, let $(u, v) \in E(D)$ and note that v is introduced at node $\phi(v)$. We claim $u \in \chi(\phi(v))$. But since ϕ is ordered, $\phi(u) < \phi(v)$ and in particular $\phi(v) \neq 0$. Then $u \in L^*_D(\phi(v) - 1, \phi) \subseteq \chi(\phi(v))$ as claimed. $\square$

4. Resolution proofs of minimal width

Recall, the ordered treewidth of a proof containing an application of the cut rule is at least 2 (Remark 3.2). Clearly, when talking about the ordered pathwidth or ordered treewidth of a proof we mean the ordered pathwidth or ordered treewidth of its underlying DAG.

4.1. Minimal ordered pathwidth

A Resolution refutation of Γ is in *Input Resolution* if it contains only applications of the cut rule and each such application has at least one premiss (i.e., label of a predecessor) in Γ (see, e.g., [26]).

Theorem 4.1. *Let ℓ be a natural and Γ a set of clauses. There is a Resolution refutation of Γ of ordered pathwidth at most 2 and length at most ℓ if and only if there is an Input Resolution refutation of Γ of length at most ℓ.*

This allows us to think of ordered pathwidth as a measure of how far a Resolution refutation is from being in Input Resolution.

To prove this we need some preparations. A digraph is *triangle-free* if so is its underlying graph. A clause is *tautological* if it contains (as terms) a variable and its negation.

Lemma 4.2. *Let $w \in \mathbb{N}$ and $(D, (F_v)_{v \in V(D)})$ be a Resolution refutation of a set Γ of clauses such that D has a width w ordered tree decomposition with underlying tree T. Then there is a Resolution refutation $(D', (F'_v)_{v \in V(D')})$ of Γ such that*

1. *$V(D') \subseteq V(D)$ contains the sink of D;*
2. *D' has an ordered tree decomposition with underlying tree T of width at most w;*

3. *no vertex in $V(D')$ has in-degree 1 in D';*
4. *no $v \in V(D')$ has a tautological label F'_v;*
5. *D' is triangle-free.*

Proof. Let Γ be a set of clauses, $(D, (F_v)_{v\in V(D)})$ a Resolution refutation of Γ, and (T, χ) an ordered tree decomposition of D of width at most w. Let v^* denote the sink of D.

In a *first step* we transform $(D, (F_v)_{v\in V(D)})$ into a Resolution refutation $(D_1, (F_v)_{v\in V(D_1)})$ where no F_v for $v \in V(D_1)$ is tautological, and where D_1 is a sub-DAG of D with unique sink v^*. If there is some $v \in V(D)$ with tautological F_v, then there is such a v having a successor w with non-tautological F_w (the sink label is not tautological). Clearly F_w must be obtained by a cut from F_v and $F_{w'}$, where w' is the other predecessor of w. Then F_w is a weakening of $F_{w'}$ and we delete the edge (v, w). The deletion of (v, w) may have caused that v has become a sink. Then we repeatedly delete sinks different from v^*. Iterating this leads to a refutation $(D_1, (F_v)_{v\in V(D_1)})$ as desired.

In a *second step* we transform $(D_1, (F_v)_{v\in V(D)})$ into a Resolution refutation $(D_2, (F'_v)_{v\in V(D_2)})$ such that $F'_v \subseteq F_v$ for all $v \in V(D_2)$ and all weakenings are *improper* in the sense that if F'_v is obtained from F'_u by weakening, then $F'_v = F'_u$. Let $\phi : V(D_1) \to [|V(D_1)|]$ be a linear layout of D_1 and write $v_i = \phi^{-1}(i)$. Define F'_{v_i} recursively for each $i \in [|V(D_1)|]$ as follows. If v_i is a source in D_1, we set $F'_{v_i} := F_{v_i}$ (in particular this is the case for $i = 0$). If F_{v_i} is obtained by weakening from F_{v_j} for $j < i$, we set $F'_{v_i} := F'_{v_j}$. If F_{v_i} is obtained by a cut from F_{v_j} and F_{v_k} for $j, k < i$, then

– either F_{v_i} is a weakening of F'_{v_j} or of F'_{v_k} and we set $F'_{v_i} := F'_{v_j}$ resp. $F'_{v_i} := F'_{v_k}$;
– or otherwise F_{v_i} is a weakening of a clause F obtainable by cut on F'_{v_j} and F'_{v_k} and we set $F'_{v_i} := F$.

The digraph D'_2 is obtained from D_1 by deleting edges (v_j, v_i) resp. (v_k, v_i) in the first case above. Then D_2 is the digraph induced in D'_2 on the ancestors of the sink v^*.

Finally, in a *third step* we obtain a DAG D' from D_2 by contracting all edges (u, v) such that F'_v is obtained by weakening from F'_u. As weakenings are improper, such contractions preserve the property of being a refutation. In fact, $(D', (F'_v)_{v\in V(D')})$ is as desired: (1), (3) and (4) are easy to see and (2) follows from Lemmas 3.4 and 3.5. We verify (5). For contradiction, assume $\underline{D'}$ contains a triangle. As D' is a acyclic, this means there are $u, w, v \in V(D')$ such that $(u, v), (w, v), (u, w) \in E(D)$. Choose literals ℓ, ℓ' from F_u such that F_v is obtained cutting F_u with F_w on ℓ and F_w is obtained cutting F_u with $F_{w'}$ on ℓ', where w' is the second predecessor of w. In particular, $\neg\ell'$ is *in* $F_{w'}$ (formally, $\{\neg\ell'\} \in F_{w'}$) and $\neg\ell$ in F_w. First note that $\neg\ell'$ is not in F_w as it is cut from $F_{w'}$, so would have to appear in F_u and then F_u would be tautological. That F_u is not tautological, clearly implies $\ell \neq \neg\ell'$. Further $\ell \neq \ell'$ because otherwise $\neg\ell' = \neg\ell$ would be in F_w. Hence, ℓ and ℓ' are literals over distinct variables. But then ℓ comes into F_w from the premiss F_u. As $\neg\ell$ is in F_w, this clause is tautological, a contradiction. □

Proof of Theorem 4.1. To see the backward direction, let $(D, (F_v)_{v\in V(D)})$ be an Input Resolution refutation of Γ. We can assume that the digraph D has vertices $V(D) = \{v_i \mid i \leq n\} \cup \{v'_i \mid i < n\}$ and edges $E(D) = \{(v_i, v_{i+1}), (v'_i, v_{i+1}) \mid i < n\}$ for some natural n. Then we have an ordered path decomposition (P, χ) of D where $V(P) := [n], E(P) := \{(i, i + 1) \mid i < n\}$ and $\chi(i) := \{v_i, v'_i, v_{i+1}\}$ for $i \in [n]$.

To verify the forward direction, let $(D, (F_v)_{v\in V(D)})$ be a refutation of Γ and let (T, χ) be a path decomposition of D of width at most 2. By Lemma 4.2 we can assume that D satisfies (3) and (5). We claim that every non-source has at least one source as a predecessor.

Assume otherwise, say $v \in V(D)$ has predecessors u_1, u_2 in D which are not sources of D. By Proposition 3.17 there exists an ordered linear layout ϕ of D of out-degree 2. As the layout is ordered $\phi(u_1), \phi(u_2) < \phi(v)$. Assume $\phi(u_1) < \phi(u_2)$ (the case $\phi(u_2) < \phi(u_1)$ is symmetrical) and consider the predecessors w_1, w_2 of u_2 in D'. We can assume $\phi(w_1) < \phi(w_2) < \phi(u_2)$. Further we have that w_1, w_2, u_1 are pairwise distinct because otherwise u_1, u_2, v would form a triangle in $\underline{D}$. Hence $\phi(w_2) < \phi(u_1)$ or $\phi(w_2) > \phi(u_1)$. In the first case, $w_1, w_2, u_1 \in L^*_D(\phi(u_1), \phi)$, so $|L^*_D(\phi(u_1), \phi)| > 2$, a contradiction. In the second case, $\phi(w_2) > \phi(u_1)$ and $w_1, w_2, u_1 \in L^*_D(\phi(w_2), \phi)$, again a contradiction. □

4.2. Minimal ordered treewidth

Recall that treelike refutations have ordered treewidth 2 (Example 3.3 (1)). We prove a weak converse to this observation. This allows us to think of ordered treewidth as a measure of how far a Resolution refutation is from being treelike.

Theorem 4.3. *Let ℓ be a natural and Γ a set of clauses. If there is a Resolution refutation of Γ of ordered treewidth at most 2 and length at most ℓ, then there is a treelike Resolution refutation of Γ of length $< 2\ell$.*

Proof. Let $(D, (F_v)_{v \in V(D)})$ be a refutation of Γ with $otw(D) = 2$. By Lemma 4.2 we can assume that D satisfies (3) and (5). Let (T, χ) be a width 2 ordered tree decomposition of D. We claim that D is *almost treelike* in the sense that all its vertices of out-degree ≥ 2 are sources.

This implies the theorem: for each source v with $r \geq 2$ successors $w_1, \ldots, w_r$ replace the edges $(v, w_2), \ldots, (v, w_r)$ by edges $(v_2, w_2), \ldots, (v_r, w_r)$ for $r - 1$ new vertices $v_2, \ldots, v_r$; label the new vertices by F_v. This gives a treelike refutation, say with underlying DAG D'. Note $|E(D)| = |E(D')|$ by construction and $|V(D')| = |E(D')| + 1$ since D' is a tree. Further, $|E(D)| \leq 2(|V(D)| - 1)$ because D has in-degree at most 2 and there is at least one source. Hence $|V(D')| \leq 2(|V(D)| - 1) + 1 < 2|V(D)|$ as claimed.

To prove our claim, we show that D_v is almost treelike for every $v \in V(D)$; here, D_v is the sub-DAG of D induced on the ancestors of v in D.

This is clear for sources v. If v is not a source, but has predecessors u_1, u_2 we assume that both D_{u_1} and D_{u_2} are almost treelike, and show that also D_v is almost treelike.

Assume $u \in V(D_v)$ has out-degree ≥ 2 in D_v. We have to show that u is a source.

Claim 1. $u \in V(D_{u_1}) \cap V(D_{u_2})$.

Proof of Claim 1. Since $V(D_v) = V(D_{u_1}) \cup V(D_{u_2}) \cup \{v\}$ and $u \neq v$ we can assume that u is in $V(D_{u_1})$. For the sake of contradiction assume $u \notin V(D_{u_2})$. As D_{u_2} is closed under predecessors, no successor of u is in $V(D_{u_2})$. But u has at least two successors w, w' and these cannot be both in $V(D_{u_1})$ because D_{u_1} is almost treelike. Hence one of them, say w, equals v, and $w' \in V(D_{u_1})$. Then u is a predecessor of v outside $V(D_{u_2})$, so $u = u_1$. It follows that w' is both a successor and an ancestor of u and this contradicts acyclicity. ⊣

If one of u_1, u_2 is a source, say u_1, then $V(D_{u_1}) = \{u_1\}$. By Claim 1 then $u = u_1$ and we are done. Hence, assume that none of u_1, u_2 is a source. Choose $t \in V(T)$ where v is introduced. Then $u_1, u_2 \in \chi(t)$ so we find ancestors t_1, t_2 of t in T where u_1, u_2 are introduced respectively.

Claim 2. t_1, t_2 are incomparable in the sense that none is an ancestor of the other.

Proof of Claim 2. Assume otherwise, say, t_2 is an ancestor of t_1. Then t_1 lies on the path in T from t_2 to t and hence $u_2 \in \chi(t_1)$. As u_1 is not a source and introduced at t_1, the bag $\chi(t_1)$ also contains the two predecessors w_1, w_2 of u_1. Then $u_1, u_2, w_1, w_2 \in \chi(t_1)$ so these vertices cannot be pairwise distinct. Then $u_2 \in \{w_1, w_2\}$. It follows that $\{u_1, u_2, v\}$ induces a triangle in $\underline{D}$, a contradiction. ⊣

So we know t_1, t_2 are incomparable, say with t_0 as least upper bound, i.e., t_0 has both t_1, t_2 as ancestors but no predecessor of t_0 has this property. This t_0 lies on the paths in $\underline{T}$ from t_1 and t_2 to t, so $u_1, u_2 \in \chi(t_0)$.

It is not hard to show that all ancestors of u_1 in D are introduced at an ancestor of t_1 in T; similarly for u_2 and t_2. In particular u is introduced at ancestors s_1, s_2 of t_1, t_2 respectively. The path in $\underline{T}$ from s_1 to s_2 contains the path from t_1 to t_2, and hence contains t_0. We finally show that u is a source. Otherwise its two predecessors are different from u_1 and from u_2. As they are in $\chi(s_1)$ and well as in $\chi(s_2)$, it follows they are in $\chi(t_0)$. As $\chi(t_0)$ also contains u_1, u_2 it has cardinality at least 4, a contradiction. □

5. Proof space

Let $k, w, \ell > 0$ be naturals, F a k-DNF and Γ a set of k-DNFs.

5.1. Ordered pathwidth is proof space

In the Introduction we informally explained a bounded space proof by a sequence of blackboards. Formally, we follow [18] and define a *space w $R(k)$-proof of F from* Γ to be a finite sequence $(\mathbb{B}_0, \ldots, \mathbb{B}_{\ell-1})$ of sets $\mathbb{B}_i$ of k-DNFs, called *blackboards*, each of cardinality at most w such that $\mathbb{B}_0 = \emptyset$ and $F \in \mathbb{B}_{\ell-1}$ and for all $0 < i < \ell$ there is a formula G such that

(B1) $\mathbb{B}_i = \mathbb{B}_{i-1} \cup \{G\}$ and $G \in \Gamma$, or

(B2) $\mathbb{B}_i = \mathbb{B}_{i-1} \cup \{G\}$ and G is derived from at most two formulas in $\mathbb{B}_{i-1}$ by one application of some inference rule of $R(k)$, or

(B3) $\mathbb{B}_i = \mathbb{B}_{i-1} \setminus \{G\}$.

The space measure above is known as "formula space" or, in case $k = 1$, as "clause space." In the Introduction we also mentioned an alternative definition of proof space (used e.g. in [5]) as a number associated with a proof written as a sequence of formulas. This definition reads like the out-degree of a linear layout. In view of Proposition 3.17 the following observation just spells out in what sense the two definitions are equivalent. Combining with Theorem 4.1 this allows to think of the space of a Resolution refutation as a measure of how far it is from being in Input Resolution.

Proposition 5.1.

1. *If there is a space w $R(k)$-proof of F from Γ of length ℓ, then there is an $R(k)$-proof of F from Γ of length $< \ell$ and ordered pathwidth $< w$.*
2. *If there is an $R(k)$-proof of F from Γ of length ℓ and ordered pathwidth w, then there is a space $(w + 1)$ $R(k)$-proof of F from Γ of length at most 2ℓ.*

Proof. (1). Let $(\mathbb{B}_0, \dots, \mathbb{B}_{\ell-1})$ be a space w $R(k)$-proof of F from Γ. We assume F appears first in $\mathbb{B}_{\ell-1}$ and consecutive blackboards are distinct. Say there are $\ell' \leq \ell - 1$ many (B1) or (B2) inferences, and say, the vth of them derives blackboard $\mathbb{B}_{i_v}$ adding the formula F_v. For a (B2) inference choose *maximal* indices $u, u' < v$ such that F_v is derived from $F_u, F_{u'}$. These pairs $(u, v), (u', v)$ are the edges of a DAG D with $V(D) = [\ell']$ and the labeling $(G_v)_{v\in[\ell']}$ makes it an $R(k)$-proof of $F = F_{\ell'}$ from Γ; the identity ϕ on $[\ell']$ is an ordered linear layout. By Proposition 3.17 it suffices to show $|L^*_D(j, \phi)| < w$ for all $j \in [\ell']$. But the formulas F_u for $u \in L^*_D(j, \phi)$ are pairwise distinct and distinct from F_{j+1} and appear in $\mathbb{B}_{i_{j+1}}$. Hence $|L^*_D(j, \phi)| \leq |\mathbb{B}_{i_{j+1}} \setminus \{F_{j+1}\}| < w$.

(2). Assume there is an $R(k)$-proof of F from Γ of length ℓ and ordered pathwidth w. By Lemmas 3.11 and 3.10, we find a width w ordered path decomposition with underlying path with $\leq \ell$ nodes. Consider the sequence of bags as they appear along the path up to the bag where the sink (labelled F) is introduced. Replace each bag by the size $\leq w + 1$ blackboard containing the labels of its vertices. If necessary, add a starting blackboard $\emptyset$ and some blackboards in between to make this a space $w + 1$ proof. This proof has a (B1) or (B2) inference whenever a vertex is introduced and a (B3) inference whenever a vertex outside the last bag is forgotten. As each vertex is forgotten and introduced exactly once, this gives $\leq \ell$ (B1) or (B2) inferences, and $\leq \ell - 1$ (B3) inferences. □

5.2. Ordered treewidth as interactive proof space

The conversation of a teacher with her student described informally in the Introduction is described more formally by a game $\Pi^k_w(\Gamma, F)$ between two players called *Student* and *Teacher* on the following game graph.

Its vertices are partitioned into *Student positions* and *Teacher positions*, the former are $R(k)$-proofs of length at most w and the latter are k-DNFs. Its directed edges run from each k-DNF to all length $\leq w$ proofs of it, and from each proof to all labels of its sources that are outside of Γ. In particular, precisely the proofs from Γ are sinks. The *initial* position is the Teacher position F. A *play* of length $0 < n \in \mathbb{N} \cup \{\infty\}$ is a sequence $(v_i)_{i<n}$ of positions with edges from v_i to v_{i+1} and v_0 being the initial position. A *strategy for Teacher* (in $\Pi^k_w(\Gamma, F)$) is a function that maps finite plays ending in a Teacher position to a successor of this position; it is *positional* in case this value depends only on the Teacher position reached by the play. A play *conforms* with the strategy if every Student position in it is the value of the strategy on the initial segment of the play up to it. The strategy is *winning* if all plays conforming with it are finite, and *ℓ-winning* if all plays conforming with it have length at most $2\ell - 1$, i.e., Teacher wins making at most ℓ moves.

Remark 5.2. The game $\Pi^k_w(\Gamma, F)$ can be seen as a parity game, so it is memory-less determined; in particular, if a winning strategy for Teacher exists, then so does a positional one [34].

By a standard argument we get the following result.

Proposition 5.3. *If there is an ℓ-winning strategy for Teacher in $\Pi^k_w(\Gamma, F)$, then there is also a positional one.*

Proof. Assume there is an ℓ-winning strategy for Teacher in $\Pi^k_w(\Gamma, F)$. Let W_i be the set of Teacher positions G such that an i-winning strategy for teacher in $\Pi^k_w(\Gamma, G)$ exists. Then W_1 is the set of predecessors of sinks, i.e., formulas that have length $\leq w$ $R(k)$-proofs from Γ. Recursively, W_{i+1} equals W_i plus those G that have an *i-good* successor, namely one all of whose successors are in W_i; in other words, W_{i+1} is the set of formulas with length $\leq w$ $R(k)$-proofs from $W_i \cup \Gamma$.

An "attractor strategy" maps plays ending in some $G \in W_1$ to a sink and plays ending in some $G \in W_{i+1} \setminus W_i$ to an i-good successor of it. Such a strategy is positional and ℓ-winning in $\Pi^k_w(\Gamma, F)$; note $F \in W_\ell$ by assumption. □

Theorem 5.4. *There is an ℓ-winning strategy for Teacher in $\Pi^k_w(\Gamma, F)$ if and only if there is an $R(k)$-proof of F from Γ with an ordered tree-decomposition of width $< w$ and height $< \ell$.*

Proof. Assume there is an ℓ-winning strategy for Teacher in $\Pi^k_w(\Gamma, F)$. By the previous proposition, we can assume the strategy is positional. Consider the following tree T with nodes t labeled with Student positions $\pi(t)$. The label of the root is the value the winning strategy gives the initial position. Every node t has exactly one predecessor (in T) for each of the at most w many successors of $\pi(t)$ in the game graph; the label of such a node is the value given by the strategy to the corresponding successor. Note that the leafs of T are labeled with sinks in the game graph, i.e., with length $\leq w$ proofs from Γ. The labels on branches of this tree correspond to the sequence of Student positions in a play conforming with the strategy. Since the strategy is ℓ-winning, T has height at most $\ell - 1$.

We can assume that the sets of vertices of (the DAGs underlying) the proofs $\pi(t)$, for $t \in V(T)$, are pairwise disjoint. For each t with successor t' in T identify the sink of $\pi(t)$, say labeled G, and all sources of $\pi(t')$ with label G. This ensures that the union of the $\pi(t), t \in V(T)$, is an $R(k)$-proof. Letting $\chi(t)$ denote the vertices of $\pi(t)$, then (T, χ) witnesses that this proof has ordered treewidth at most $w - 1$.

Conversely, let $(D, (F_v)_{v \in V(D)})$ be an $R(k)$-proof of F from Γ and (T, χ) an ordered tree decomposition of D of height $< \ell$ and width $< w$. We can assume that $\chi(t) \neq \emptyset$ for all $t \in V(T)$. For $t \in V(T)$ and $v \in \chi(t)$ let $\pi(t, v)$ be the proof induced on the ancestors of v in $D[\chi(t)]$. We informally describe a winning strategy for Teacher.

On the initial position F, Teacher chooses $v_1 \in V(D)$ and $t_1 \in V(T)$ such that $F_{v_1} = F$ and v_1 is introduced at t_1, and moves to $\pi(t_1, v_1)$. If Student moves to source label G, Teacher chooses a source v_2 of $\pi(t_1, v_1)$ such that $G = F_{v_2}$, chooses an ancestor t_2 of t_1 where v_2 is introduced and answers $\pi(t_2, v_2)$. And so on. Note that a strategy implementing such moves is not positional. Namely, for her ith move Teacher remembers a vertex $v_i \in V(D)$ and a node $t_i \in V(T)$ and these satisfy

(a) F_{v_i} is a Teacher position in the play,
(b) v_i is introduced at t_i,
(c) v_{i+1} is a source in $\pi(t_i, v_i)$,
(d) t_{i+1} is an ancestor of t_i.

Note, no Teacher position in the play is a formula in Γ (here we assume $F \notin \Gamma$; otherwise there is a 1-winning strategy). In particular, $F_{v_{i+1}} \notin \Gamma$ (by (a)), so v_{i+1} has predecessors in D. These are not in $\pi(t_i, v_i)$ (by (c)) and, by definition of $\pi(t_i, v_i)$, also not in $\chi(t_i)$. As the tree-decomposition is ordered, v_{i+1} is not introduced at t_i. As v_{i+1} is introduced at t_{i+1} (by (b)) we have $t_{i+1} \neq t_i$. By (d) then the sequence $t_1, t_2, \ldots$ has length $\leq \ell$. This implies that the strategy is ℓ-winning. □

Remark 5.5. Assume Γ is a set of clauses. If Teacher wins $\Pi^k_w(\Gamma, 0)$, then Γ is contradictory and hence has a treelike Resolution refutation. This refutation has ordered treewidth at most 2, so by Theorem 5.4 Teacher wins $\Pi^1_3(\Gamma, 0)$. It follows that for $w \geq 3$, Teacher wins $\Pi^k_w(\Gamma, 0)$ if and only if Teacher wins $\Pi^1_3(\Gamma, 0)$. Thus, the parameters k and w only matter when taking into account how fast Teacher can win, that is, when considering ℓ-winning strategies.

As a side remark we observe that if the teacher knows how to convince visiting students very quickly then she also does not need a large blackboard in class.

Corollary 5.6. *If Teacher has an ℓ-winning strategy in $\Pi^k_w(\Gamma, F)$, then there is a space ℓw $R(k)$-proof of F from Γ.*

Proof. Assume Teacher has an ℓ-winning strategy in $\Pi^k_w(\Gamma, F)$. By the previous result there is an $R(k)$-proof of F from Γ with an ordered tree decomposition (T, χ) of width $\leq w - 1$ and height $\leq \ell - 1$. By Proposition 3.13 this proof has ordered pathwidth $< w \cdot \ell$. Now apply Theorem 5.1 (2). □

Remark 5.7. It is well-known and easy to see that every contradictory set of clauses Γ has a treelike Resolution refutation of height at most the number n of variables in Γ. By the previous remark this gives an n-winning strategy of Teacher in $\Pi^1_3(\Gamma, 0)$ and the last corollary thus shows that Γ has a Resolution refutation in space $3n$. Even $n+1$ is known and proved in [18, Theorem 12].

6. Lower bounds

Recall Definition 2.5.

Theorem 6.1. *Let φ be a first-order τ-sentence ready for translation that has an infinite model. Let r be the maximal arity of some function symbol in τ and assume $r \geq 1$.*

Then there exists a real $c_\varphi > 0$ such that for every natural $n \geq 1$ and every natural $k \geq 1$, every strongly sound k-DNF refutation $(D, (F_v)_{v \in V(D)})$ of $\langle\varphi\rangle_n$ satisfies

$$k \cdot otw(D) \cdot \log |V(D)| > c_\varphi \cdot n^{1/r}.$$

Remark 6.2. The assumption that $r \geq 1$ does not exclude interesting cases. If $r = 0$, all function symbols of τ are constants. In an infinite model of φ every nonempty set containing their interpretations carries a submodel which too is a model of φ (being universal). Hence, the spectrum of φ is co-finite, so all but finitely many $\langle\varphi\rangle_n$ are satisfiable and have no sound refutations at all.

Proof. The proof carries out the sketch given in Section 1.5. Let a τ-sentence φ and a natural r accord the assumption of the theorem, and let M be an infinite model of φ. Write $\varphi = \forall \bar{x} \bigwedge_{i \in I} C_i(\bar{x})$ according Definition 2.5. Let $m_0, \ldots, m_{\ell-1}$ list without repetition the interpretations of constants of τ in M. It suffices to find $c_\varphi > 0$ satisfying our claim for every positive $n \geq \ell$.

So let $n \geq \ell$ and k be positive naturals, and let $(D, (F_v)_{v \in V(D)})$ be a strongly sound k-DNF refutation of $\langle\varphi\rangle_n$. Write $w := otw(D)$. By Lemma 3.12, D has an ordered tree decomposition of width w with a tree of in-degree at most 2 and at most $2|V(D)|$ many nodes. We can assume that all its bags are nonempty. Add the sink of D to all bags on nodes on the path from the node where it is forgotten to the root. The resulting tree decomposition (T_0, χ) has width at most $w+1$ with the sink of D contained in $\chi(r_{T_0})$, the bag at the root.

For $X \subseteq V(T_0)$ we write

$$\chi(X) := \bigcup_{t \in X} \chi(t).$$

Conditions. For $N \subseteq M$ let $\partial N := \bigcup_f \mathrm{im}(f^M \upharpoonright N)$, where f ranges over the function symbols of τ, i.e., ∂N contains the values which M's functions take on N. Note $m_0, \ldots, m_{\ell-1} \in \partial N$ for every $N \subseteq M$ and

$$|\partial N| \leq |\tau| \cdot \max\{|N|, 1\}^r. \tag{2}$$

We define a *condition* to be a pair (κ, λ) of partial bijections from $[n]$ to M such that $\kappa \subseteq \lambda$ and $\mathrm{im}(\lambda) = \mathrm{im}(\kappa) \cup \partial \mathrm{im}(\kappa)$. We say a condition (κ^*, λ^*) *extends* another (κ, λ) if $\kappa \subseteq \kappa^*$ and $\lambda \subseteq \lambda^*$. With a condition (κ, λ) we associate the restriction $\rho(\kappa, \lambda)$ which is defined on a propositional atom of the form $R\bar{a}$ or $f\bar{a} = b$ if and only if κ is defined on all components of $\bar{a}$; in this case it maps

- $R\bar{a}$ to 1 if $\kappa(\bar{a}) \in R^M$, and to 0 otherwise;
- $f\bar{a}{=}b$ to 1 if $\lambda^{-1}(f^M(\kappa(\bar{a}))) = b$, and to 0 otherwise;

note that λ^{-1} is defined on $f^M(\kappa(\bar{a})) \in \partial \mathrm{im}(\kappa)$. Here, for a tuple $\bar{a} = a_1 \cdots a_s$, we write $\kappa(\bar{a})$ for the tuple $\kappa(a_1) \cdots \kappa(a_s)$.

Observe that, if (κ^*, λ^*) extends (κ, λ) in the sense above, then $\rho(\kappa^*, \lambda^*)$ extends $\rho(\kappa, \lambda)$ as a partial function. The *rank* of (κ, λ) is $|\mathrm{dom}(\kappa)|$. For example, $(\emptyset, \lambda_0)$ is a condition of rank 0, where λ_0 is the function that maps $i < \ell$ to m_i.

Claim 1. If (κ, λ) is a condition and C a clause in $\langle\varphi\rangle_n$, then $C \upharpoonright \rho(\kappa, \lambda) \neq 0$.

Proof of Claim 1. We assume $\rho(\kappa, \lambda)$ is defined on all variables appearing in C (otherwise there is nothing to show). If C is a functionality clause, either $\bigvee_b f\bar{a}{=}b$ or $\neg f\bar{a}{=}b \vee \neg f\bar{a}{=}b'$, then $\rho(\kappa, \lambda)$ is defined on all variables of the form $f\bar{a}{=}c$ for $c \in [n]$. By definition it evaluates exactly one of them, namely the one for $c := \lambda^{-1}(f^M(\kappa(\bar{a})))$, to 1 and all others to 0. This implies $C \upharpoonright \rho(\kappa, \lambda) = 1$.

Suppose C is $\langle C_i(\bar{a})\rangle$ and choose an injection $\lambda' : [n] \to M$ extending λ (we only need it to be defined on all components of $\bar{a}$). If $\langle C_i(\bar{a})\rangle = 1$, there is nothing to show. Otherwise all equality literals in $C_i(\bar{x})$ become "false" under the substitution of $\bar{a}$ for $\bar{x}$. Since λ' is injective, the tuple $\lambda'(\bar{a})$ falsifies all these literals in M. But since M is a model of φ, the tuple $\lambda'(\bar{a})$ satisfies $C_i(\bar{x})$ in M, so satisfies some literal mentioning a symbol from τ. Writing $\bar{x} = x_0 \cdots x_{s-1}$ and $\bar{a} = a_0 \cdots a_{s-1}$ we can write our literal as $(\neg) f x_{i_0} \cdots x_{i_{r-1}} {=} x_{i_r}$ or $(\neg) R x_{i_0} \cdots x_{i_{r-1}}$ where $f, R \in \tau$ are r-ary symbols for some $r \in \mathbb{N}$ and $i_0, \ldots, i_r \in [s]$. Assume our literal is $f x_{i_0} \cdots x_{i_{r-1}} {=} x_{i_r}$, the other cases are treated analogously. Then $f^M(\lambda'(a_{i_0}), \ldots, \lambda'(a_{i_{r-1}})) = \lambda'(a_{i_r})$ and $f a_{i_0} \cdots a_{i_{r-1}} {=} a_{i_r}$ is a propositional literal in $\langle C_i(\bar{a})\rangle$. Since we assumed $\rho(\kappa, \lambda)$ to be defined on all atoms in $\langle C_i(\bar{a})\rangle$, we have that $a_{i_0}, \ldots, a_{i_{r-1}} \in \mathrm{dom}(\kappa)$ and $f^M(\kappa(a_{i_0}), \ldots, \kappa(a_{i_{r-1}})) = \lambda'(a_{i_r})$. Hence $\lambda'(a_{i_r}) \in \mathrm{im}(\lambda)$, and as $\lambda' \supseteq \lambda$ is an injection, $\lambda^{-1}(\lambda'(a_{i_r})) = a_{i_r}$. By definition then the restriction $\rho(\kappa, \lambda)$ evaluates $f a_{i_0} \cdots a_{i_{r-1}} {=} a_{i_r}$ to 1 and $C \restriction \rho(\kappa, \lambda) = 1$ follows. ⊣

Claim 2. Let $B \subseteq [n]$ and assume (κ, λ) is a condition of rank at most d. If

$$n \geq 3|\tau| \cdot (d + |B|)^r, \tag{3}$$

then there exists a condition (κ', λ') extending (κ, λ) such that $B \subseteq \mathrm{dom}(\kappa')$.

Proof of Claim 2. We can assume that $B \neq \emptyset$. Let $\tilde{d} := \max\{d, 1\}$. By (2) we have

$$|\mathrm{dom}(\lambda)| \leq |\partial\mathrm{im}(\kappa)| + |\mathrm{dom}(\kappa)| \leq |\tau| \cdot \tilde{d}^r + d.$$

Choose some minimal injective extension κ' of κ which is compatible with λ (i.e. $\kappa' \cup \lambda$ is a function) and such that $B \subseteq \mathrm{dom}(\kappa')$. This is possible if $n - \mathrm{dom}(\lambda) \geq |B|$ and hence if $n \geq (|\tau| \cdot \tilde{d}^r + d) + |B|$. This follows from (3) (note $|\tau| > 0$ as $r \geq 1$).

Then choose a minimal injective extension λ' of $\kappa' \cup \lambda$ such that $\mathrm{im}(\lambda') \supseteq \partial\mathrm{im}(\kappa')$. By (2) we have $|\partial\mathrm{im}(\kappa')| \leq |\tau| \cdot (|B| + d)^r$. Hence, the choice of λ' is possible if $n - |\mathrm{dom}(\kappa') \cup \mathrm{dom}(\lambda)|$ is at least $|\tau| \cdot (|B| + d)^r$. Since $|\mathrm{dom}(\kappa') \cup \mathrm{dom}(\lambda)| \leq |B| + |\mathrm{dom}(\lambda)|$, the choice of λ' is possible if $n \geq |\tau| \cdot (d + |B|)^r + |B| + (|\tau| \cdot \tilde{d}^r + d)$. This is also implied by (3). ⊣

Adversary positions. An *adversary position* is a tuple (T, L, κ, λ) such that

(A1) T is a subtree of T_0 which is fully in T_0 (cf. Definition 3.7);
(A2) $L \subseteq V(T_0)$ contains every leaf of T which is not a leaf of T_0;
(A3) (κ, λ) is a condition such that
  (A3a) $F_v \restriction \rho(\kappa, \lambda) = 1$ for all $v \in \chi(L)$, and
  (A3b) for every condition (κ^*, λ^*) such that (κ^*, λ^*) extends (κ, λ) there exists $v \in \chi(r_T)$ such that $F_v \restriction \rho(\kappa^*, \lambda^*) \neq 1$.

Adversary positions exist: for example, $(T_0, \emptyset, \emptyset, \lambda_0)$ is one; property (A3b) holds because $\chi(r_{T_0})$ contains the sink v of D and $F_v = 0$.

Claim 3. Suppose (T, L, κ, λ) is an adversary position and $v \in \chi(r_T)$. Let

$$\Gamma_T := \{F_u \mid u \in \chi(V(T)) \text{ is a source in } D\}.$$

Then there exists a strongly sound k-DNF proof of F_v from $\Gamma_T \cup \{F_u \mid u \in \chi(L)\}$.

Proof of Claim 3. Let $D' := D[\chi(V(T))]$. Observe $V(T) \neq \emptyset$ as T is a subtree, so $\chi(V(T)) \neq \emptyset$ as bags are nonempty. It suffices to show that for every $v \in V(D')$ either all predecessors of v in D are in $V(D')$ or $v \in \chi(L)$. But if $(u, v) \in E(D)$ and $u \notin V(D')$, then Lemma 3.8 (2) and (A1) imply that $v \in \chi(t)$ for some leaf t of T which is not a leaf in T_0; by (A2) then $v \in \chi(L)$. ⊣

Recall that T_t denotes the subtree of a tree T rooted at t. An adversary position (T, L, κ, λ) has as *successor* any tuple $(T', L', \kappa', \lambda')$ that can be obtained as follows.

> Choose $t \in V(T)$ such that both $V(T_t)$ and $V(T) \setminus (V(T_t) \setminus \{t\})$ have cardinality at most $\lfloor 2|V(T)|/3 \rfloor + 1$. Such a t exists because T has in-degree at most 2 (as a subtree of T_0).
>
> *Case 1.* Property (A3b) holds for t, i.e. for every extension (κ^*, λ^*) of (κ, λ) there exists $v \in \chi(t)$ such that $F_v \restriction \rho(\kappa^*, \lambda^*) \neq 1$.

Set $T' := T_t, L' := L, \kappa' := \kappa$ and $\lambda' := \lambda$.

Case 2. Otherwise, choose an extension (κ^*, λ^*) of (κ, λ) of minimal rank among those satisfying $F_v \upharpoonright \rho(\kappa^*, \lambda^*) = 1$ for all $v \in \chi(t)$.

Set $T' := T - (V(T_t) \setminus \{t\}), L' := L \cup \{t\}, \kappa' := \kappa^*$ and $\lambda' := \lambda^*$.

Trivially, every adversary position has successors.

Claim 4. If (T, L, κ, λ) is an adversary position with successor $(T', L', \kappa', \lambda')$, then $(T', L', \kappa', \lambda')$ too is an adversary position. The rank of (κ', λ') is at most $(w+2) \cdot k \cdot r_\varphi$ bigger than the rank of (κ, λ) where r_φ denotes the maximal arity of some symbol in τ.

For the proof, we use a mode of speech from [16] and say that the propositional variables $R\bar{a}$ and $f\bar{a} = b$ *mention* an element $a \in [n]$ if a is a component of $\bar{a}$; in particular $f\bar{a} = b$ does not necessarily mention b. A formula mentions an element if so does some variable appearing in it.

Proof of Claim 4. Let $t \in V(T)$ be the node chosen to compute $(T', L', \kappa', \lambda')$ from (T, L, κ, λ). Both subtrees T_t and $T - (V(T_t) \setminus \{t\})$ are fully in T. Since T is fully in T_0, so is T' and $(T', L', \kappa', \lambda')$ satisfies (A1). Properties (A2) and (A3a) are clear. Property (A3b) follows in Case 1 because $r_{T'} = t$, and in Case 2 because $r_{T'} = r_T$, (κ', λ') extends (κ, λ) and (T, L, κ, λ) satisfies (A3b).

To see the second statement, assume $(T', L', \kappa', \lambda')$ is obtained according to Case 2 (in Case 1 there is nothing to show). Choose a condition $(\tilde{\kappa}, \tilde{\lambda})$ extending (κ, λ) such that $F_v \upharpoonright \rho(\tilde{\kappa}, \tilde{\lambda}) = 1$ for every $v \in \chi(t)$. For $v \in \chi(t)$ choose a k-term t_v in the k-DNF F_v such that $t_v \upharpoonright \rho(\tilde{\kappa}, \tilde{\lambda}) = 1$. Any restriction ρ that agrees with $\rho(\tilde{\kappa}, \tilde{\lambda})$ on the atoms appearing in these k-terms is such that $F_v \upharpoonright \rho = 1$ for every $v \in \chi(t)$. In particular, this is the case for $\rho(\tilde{\kappa} \restriction (A \cup \mathrm{dom}(\kappa)), \tilde{\lambda}')$ where A is the set of elements mentioned by $\bigwedge_{v \in \chi(t)} t_v$ and $\tilde{\lambda}'$ is a suitable restriction of $\tilde{\lambda}$ such that $(\tilde{\kappa} \restriction (A \cup \mathrm{dom}(\kappa)), \tilde{\lambda}')$ is a condition. Every k-term t_v mentions at most $k \cdot r_\varphi$ many elements, and there are at most $|\chi(t)| \leq w+2$ many terms t_v. Thus, the rank of $(\tilde{\kappa} \restriction (A \cup \mathrm{dom}(\kappa)), \tilde{\lambda}')$ and hence of (κ', λ') is at most $|A| \leq (w+2) \cdot k \cdot r_\varphi$ bigger than the rank of (κ, λ). ⊣

Wrapping up. Let $((T_i, L_i, \kappa_i, \lambda_i))_{i \in \mathbb{N}}$ be a sequence such that $(T_{i+1}, L_{i+1}, \kappa_{i+1}, \lambda_{i+1})$ is a successor of $(T_i, L_i, \kappa_i, \lambda_i)$ for all $i \in \mathbb{N}$, and $(T_0, L_0, \kappa_0, \lambda_0)$ is $(T_0, \emptyset, \emptyset, \lambda_0)$; we already noted that this is an adversary position. By Claim 4 all tuples $(T_i, L_i, \kappa_i, \lambda_i)$ are adversary positions. Further, $|V(T_{i+1})| \leq \lfloor 2|V(T_i)|/3 \rfloor + 1$, so $|V(T_m)| \leq 3$ for $m := \lceil \log_{3/2} |V(T_0)| \rceil$. Recalling that $|V(T_0)| \leq 2|V(D)|$ the theorem follows once we show

$$n < 3|\tau| \cdot (m \cdot (w+2) \cdot k \cdot r_\varphi + 3(w+2) \cdot w_\varphi \cdot r_\varphi)^r, \tag{4}$$

where w_φ is the maximal number of literals in some first order clause $C_i(\bar{x})$ of φ. Note functionality clauses mention at most $r \leq r_\varphi$ many elements of $[n]$, hence every clause in $\langle \varphi \rangle_n$ mentions at most $w_\varphi \cdot r_\varphi$ many elements of $[n]$.

We now verify (4). Since (κ_0, λ_0) has rank 0, Claim 4 implies that (κ_m, λ_m) has rank at most

$$d_m := m \cdot (w+2) \cdot k \cdot r_\varphi.$$

Recall the notation Γ_{T_m} from Claim 3 and let $B \subseteq [n]$ denote the set of elements mentioned by formulas in Γ_{T_m}. Note $\Gamma_{T_m} \subseteq \langle \varphi \rangle_n$. Since $|V(T_m)| \leq 3$ we have $|\Gamma_{T_m}| \leq 3(w+2)$ and hence

$$|B| \leq 3(w+2) \cdot w_\varphi \cdot r_\varphi.$$

Assume for contradiction that (4) fails. Then $n \geq 3|\tau| \cdot (d_m + |B|)^r$. By Claim 2 there exists a condition (κ, λ) extending (κ_m, λ_m) such that $B \subseteq \mathrm{dom}(\kappa)$. By (A3b) there exists $v_m \in \chi(r_{T_m})$ such that $F_{v_m} \upharpoonright \rho(\kappa, \lambda) \neq 1$. By Claim 3 there is a strongly sound k-DNF proof of F_{v_m} from $\Gamma_{T_m} \cup \{F_u \mid u \in \chi(L_m)\}$. For every clause $C \in \Gamma_{T_m} \subseteq \langle \varphi \rangle_n$, we have that $\mathrm{dom}(\kappa)$ contains all elements mentioned by C. Hence $\rho(\kappa, \lambda)$ evaluates every atom in C, so $C \upharpoonright \rho(\kappa, \lambda) \in \{0, 1\}$, and hence $C \upharpoonright \rho(\kappa, \lambda) = 1$ by Claim 1. Further we have $F_u \upharpoonright \rho(\kappa, \lambda) = 1$ for every $u \in \chi(L_m)$ since $\rho(\kappa, \lambda)$ extends $\rho(\kappa_m, \lambda_m)$ and (A3a). In summary, F_{v_m} does not restrict to 1 under $\rho(\kappa, \lambda)$ and there is a strongly sound k-DNF proof of F_{v_m} from formulas that do restrict to 1 under $\rho(\kappa, \lambda)$. This contradicts Lemma 2.3. □

This proof has the following corollary.

Corollary 6.3. *Let φ and r be as in Theorem 6.1. Then there exists a real $c_\varphi > 0$ such that for every natural $n \geq 1$ and every natural $k \geq 1$, every strongly sound k-DNF refutation $(D, (F_v)_{v \in V(D)})$ of $\langle \varphi \rangle_n$ satisfies*

$$k \cdot opw(D) > c_\varphi \cdot n^{1/r}.$$

Proof. Let $(D, (F_v)_{v \in V(D)})$ be a strongly sound k-DNF refutation of $\langle \varphi \rangle_n$ with path decomposition (T, χ) of width $w := opw(D)$. Assume the nodes of the path T are $0, 1, 2 \ldots, \ell$ in order. We can assume that for each i there is v_i such that $\chi(i) \setminus \{v_i\} = \chi(i+1)$ or $\chi(i) \cup \{v_i\} = \chi(i+1)$. Furthermore we assume $\chi(0) = \emptyset$.

Call a condition (κ, λ) *good for* i if for every $v \in \chi(i)$

(a) if F_v is a clause from $\langle \varphi \rangle_n$, then $\mathrm{dom}(\kappa)$ contains all elements mentioned by F_v;
(b) otherwise there is a term t_v of F_v such that $\mathrm{dom}(\kappa)$ contains all elements mentioned by t_v and $t_v \restriction \rho(\kappa, \lambda) = 1$.

Recall the constants r_φ, w_φ from the previous proof.

Claim 5. Let $i \leq \ell$. If there is a condition good for i, then there is one of rank at most

$$(w+1) \cdot k \cdot w_\varphi \cdot r_\varphi.$$

Proof of Claim 5. Let (κ, λ) be good for i. For $v \in \chi(i)$ let $B_v \subseteq [n]$ be the set of elements mentioned by F_v if $F_v \in \langle \varphi \rangle_n$, and otherwise the set of elements mentioned by t_v (chosen according (b) above). In the first case $|B_v| \leq w_\varphi \cdot r_\varphi$ and in the second $|B_v| \leq k \cdot r_\varphi$. Set $B := \bigcup_{v \in \chi(i)} B_v$ and note $|B| \leq (w+1) \cdot k \cdot r_\varphi \cdot w_\varphi$. Define $\kappa' := \kappa \upharpoonright B$ and $\lambda' := \lambda \upharpoonright (\mathrm{im}(\kappa') \cup \partial \mathrm{im}(\kappa'))$. Then (κ', λ') is good for i and has rank at most $|B|$. ⊣

Observe that if (κ, λ) is good for i, then $F_v \restriction \rho(\kappa, \lambda) = 1$ for all $v \in \chi(i)$ (in case (a) this follows from Claim 1). In particular, there is a condition good for 0 (namely $(\emptyset, \lambda_0)$ from the previous proof) but there is no condition good for i^* where $i^* \leq \ell$ is the node where the sink of D is introduced. Hence there exists $i_* < i^*$ such that there exists a condition (κ_*, λ_*) good for i_* and such that there does not exist a condition good for $i_* + 1 \leq \ell$. In particular, (κ_*, λ_*) is not good for $i_* + 1$. It follows that $\chi(i_* + 1) = \chi(i_*) \cup \{v_{i_*}\}$ and v_{i_*} is introduced at $i_* + 1$.

By Claim 5 we can assume that (κ_*, λ_*) has rank at most $d_* := (w+1) \cdot k \cdot w_\varphi \cdot r_\varphi$.

Claim 6. v_{i_*} is a source of D.

Proof of Claim 6. Otherwise, because (T, χ) is ordered, the predecessors u, w of v_{i_*} in D are present in $\chi(i_* + 1) = \chi(i_*) \cup \{v_{i_*}\}$, so $u, w \in \chi(v_{i_*})$. Since (κ_*, λ_*) is good for i_* we have $F_u \restriction \rho(\kappa_*, \lambda_*) = F_w \restriction \rho(\kappa_*, \lambda_*) = 1$. By strong soundness $F_{v_{i_*}} \restriction \rho(\kappa_*, \lambda_*) = 1$, so (b) is satisfied for v_{i_*}. Hence, (κ_*, λ_*) is good for $i_* + 1$, a contradiction. ⊣

By Claim 6, $F_{v_{i_*}}$ is a clause from $\langle \varphi \rangle_n$. Let B_* denote the set of elements mentioned by $F_{v_{i_*}}$. Any condition (κ, λ) extending (κ_*, λ_*) with $B_* \subseteq \mathrm{dom}(\kappa)$ would satisfy $F_{v_{i_*}} \restriction \rho(\kappa^*, \lambda^*) = 1$ by Claim 1 and would thus be good for $i_* + 1$. That such a condition does not exist, implies by Claim 2 that $n < 3|\tau| \cdot (d_* + |B_*|)^r$. Noting $|B_*| \leq w_\varphi \cdot r_\varphi$, the corollary follows. □

Remark 6.4. It is well-known that Input Resolution is not refutation-complete (cf. [26]). Indeed, if φ is as above, then for sufficiently large n there is no Input Resolution refutation of $\langle \varphi \rangle_n$. This follows from the above corollary and Theorem 4.1.

The above corollary generalizes bounds on space (recall Proposition 5.1) known for particular infinity axioms (cf. Introduction). Concerning the more peculiar notion of space from Section 5.2 we find it worthwhile to explicitly note the following rather direct corollary.

Corollary 6.5. *Let φ, r and c_φ be as in Corollary 6.3 and $n, k, w, \ell > 0$ be naturals. If there exists an ℓ-winning strategy for Teacher in $\Pi^k_w(\langle \varphi \rangle_n)$, then*

$$k \cdot w \cdot \ell > c_\varphi \cdot n^{1/r}.$$

Proof. Assume there exists an ℓ-winning strategy for Teacher in $\Pi^k_w(\langle \varphi \rangle_n)$. By Corollary 5.6 there exists a space $\ell \cdot w$ $R(k)$-refutation of $\langle \varphi \rangle_n$. By Proposition 5.1 (1) there exists an $R(k)$-refutation of $\langle \varphi \rangle_n$ of ordered pathwidth $< \ell \cdot w$. Corollary 6.3 implies that $k \cdot w \cdot \ell > c_\varphi \cdot n^{1/r}$. □

To finish we take up some of the issues mentioned in the Introduction spelling out some further direct consequences of Theorem 6.1.

We note that short DAG-like refutations of translations of infinity axioms need to use a large degree of parallelism in the sense that they require large ordered treewidth. This is the first statement in the corollary below. The second can be seen as a generalization of known lower bounds for treelike $R(\log)$ [29].

Corollary 6.6. *Let φ be as in Theorem 6.1.*

1. $R(\log)$*-refutations of $\langle\varphi\rangle_n$ of length at most $2^{n^{o(1)}}$ have ordered treewidth at least $n^{\Omega(1)}$.*
2. $R(\log)$*-refutations of $\langle\varphi\rangle_n$ of ordered treewidth at most $n^{o(1)}$ have length at least $2^{n^{\Omega(1)}}$.*

As mentioned in the Introduction, the infinity axiom *lnp* from Example 2.6 has polynomial length Resolution refutations [47]. The infinity axiom *wphp* from Example 2.7 has quasipolynomial length $R(\log)$-refutations [32]. Specifically for these examples we can say the following.

Corollary 6.7.

1. *Polynomial length $R(100)$-refutations of $\langle lnp\rangle_n$ have ordered treewidth $\Omega(n/\log n)$.*
2. *Quasipolynomial length $R(\log)$-refutations of $\langle wphp\rangle_n$ have ordered treewidth $\Omega(n^{0.4})$.*

Finally, we state a trade-off for length and parallelism.

Corollary 6.8. *Let $\varepsilon > 0$.*

1. *There are Resolution refutations of $\langle lnp\rangle_n$ of ordered treewidth $O(n)$ and length $n^{O(1)}$.*
2. *There are Resolution refutations of $\langle lnp\rangle_n$ of ordered treewidth 2.*
3. $R(\log)$*-refutations of $\langle lnp\rangle_n$ with ordered treewidth $O(n^{1-\varepsilon})$ have length $2^{n^{\Omega(1)}}$.*

Proof. (2) follows since treelike refutations have ordered treewidth 2, and (3) follows from Theorem 6.1. For (1) we give length $n^{O(1)}$ space $O(n)$ Resolution refutations of $\langle lnp\rangle_n$ (Proposition 5.1 (1)). This follows familiar lines, see e.g. [11, Theorem 3.1]. Consider the clauses

$$C_k^a := \bigvee_{b\in[k]} b{<}a.$$

for $k \in [n]$ and $a \in [k]$. We derive blackboards $\mathbb{B}_n, \mathbb{B}_{n-1}, \ldots, \mathbb{B}_1$ where

$$\mathbb{B}_k := \{C_k^0, \ldots, C_k^{k-1}\}.$$

Then $\mathbb{B}_1$ contains C_1^0 and a cut with axiom $\neg 0{<}0$ yields the empty clause. By an *axiom* we mean a clause in $\langle lnp\rangle_n$. The length bound will be clear, we pay attention to space.

To derive $\mathbb{B}_n$ in space $O(n)$, note each C_n^a is derivable in space $O(1)$: cut the functionality axiom $\bigvee_{b\in[n]} fa{=}b$ with the axioms $\neg fa{=}b \vee b{<}a$ for all $b \in [n]$. To derive $\mathbb{B}_k$ from $\mathbb{B}_{k+1}$ in space $O(n)$, derive each C_k^a for $a \in [k]$ in space $O(n)$ from C_{k+1}^a and C_{k+1}^k: first derive clauses $C_k^a \vee \neg b{<}k$ for each $b \in [k]$ cutting C_{k+1}^a on $k{<}a$ with axiom $\neg b{<}k \vee \neg k{<}a \vee b{<}a$; then cut all these clauses with C_{k+1}^k to get $C_k^a \vee k{<}k$; then cut with axiom $\neg k{<}k$. □

Acknowledgements

The referees helped to significantly improve the presentation of this work. The restrictions used in the proof of Theorem 6.1 come from unpublished work of Albert Atserias, Sergi Oliva and the first author. The second author thanks the ERC (European Research Council) for its support through Project COMPLEX REASON 239962.

References

[1] M. Alekhnovich, E. Ben-Sasson, A. A. Razborov, and A. Wigderson. Space complexity in propositional calculus. *SIAM Journal on Computing* 31(4):1184–1211, 2002.

[2] A. Atserias and V. Dalmau. A combinatorial characterization of resolution width. *Journal of Computer and System Sciences* 74(3):323–334, 2008.

[3] A. Atserias and M. Müller. Partially definable forcing and bounded arithmetic. *Archive for Mathematical Logic* 54(1):1–33, 2015.

[4] J. Bang-Jensen and G. Gutin. *Digraphs*. Springer Monographs in Mathematics. Springer London Ltd., London, second edition, 2009.

[5] P. Beame, C. Beck and R. Impagliazzo. Time space tradeoffs in resolution: Superpolynomial lower bounds for superlinear space. 44th Symposium on Theory of Computing, pages 213–232, 2012.

[6] E. Ben-Sasson. Size-space tradeoffs for resolution. *SIAM Journal on Computing* 38(6):2511–2525, 2009.

[7] E. Ben-Sasson and J. Nordström. Understanding space in proof complexity: Separations and trade-offs via substitutions. 2nd Symposium on Innovations in Computer Science, pages 401–416, 2011.

[8] E. Ben-Sasson, A. Wigderson. Short proofs are narrow - resolution made simple. *Journal of the ACM* 48(2):149–169, 2001.

[9] A. Beckmann, P. Pudlák and N. Thapen. Parity games and propositional proofs. *ACM Transactions on Computational Logic* 15(2): article 17, 2014.

[10] H. L. Bodlaender. A partial *k*-arboretum of graphs with bounded treewidth. *Theoretical Computer Science* 209(1-2):1–45, 1998.

[11] M. L. Bonet, N. Galesi. Optimality of size-width tradeoffs for resolution. *Computational Complexity* 10(4):261–276, 2001.

[12] S. A. Cook and R. A. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic* 44(1):36–50, Mar. 1979.

[13] S. A. Cook and P. Nguyen. *Logical Foundations of Proof Complexity*, Cambridge University Press, 2010.

[14] S. Dantchev and B. Martin. The limits of tractability in Resolution-based propositional proof systems. *Annals of Pure and Applied Logic* 163(6):656–668, 2012.

[15] S. Dantchev and B. Martin. Relativization makes contradictions harder for Resolution. *Annals of Pure and Applied Logic* 165(3): 837–857, 2014.

[16] S. Dantchev and S. Riis. On relativisation and complexity gap for resolution-based proof systems. 17th Computer Science Logic, Lecture Notes in Computer Science 2803, pages 142–154. Springer, 2003.

[17] J. L. Esteban, N. Galesi, and J. Messner. On the complexity of resolution with bounded conjunctions. *Theoretical Computer Science* 321(2-3):347–370, 2004.

[18] J. L. Esteban and J. Torán. Space bounds for resolution. *Information and Computation* 171(1):84–97, 2001.

[19] Y. Filmus, M. Lauria, M. Mikša, J. Nordström, and M. Vinyals. From small space to small width in resolution. *ACM Transactions on Computational Logic* 16(4): Article 28, 2015.

[20] J. Flum and M. Grohe. *Parameterized Complexity Theory.* Springer, 2006.

[21] R. Ganian, P. Hliněný, J. Kneis, A. Langer, J. Obdržálek and P. Rossmanith. On digraph width measures in parameterized algorithmics. 4th International Workshop on Parameterized and Exact Computation. Lecture Notes in Computer Science 5917, pages 185-197, 2009.

[22] M. Grohe. The complexity of homomorphism and constraint satisfaction problems seen from the other side. *Journal of the ACM* 54(1): Article No. 1, 2007.

[23] T. Johnson, N. Robertson, P. D. Seymour, and R. Thomas. Directed tree-width. *Journal of Combinatorial Theory Series B* 82(1):138–154, 2001.

[24] N. G. Kinnersley. The vertex separation number of a graph equals its path-width. *Information Processing Letters* 42(6):345–350, 1992.

[25] M. Kirousis and C. H. Papadimitriou. Searching and Pebbling. *Theoretical Computer Science* 47:205–218, 1986.

[26] H. Kleine Büning and T. Lettman. *Propositional logic: deduction and algorithms.* Cambridge University Press, Cambridge, 1999.

[27] J. Krajíček. Lower bounds to the size of constant-depth propositional proofs. *Journal of Symbolic Logic* 59(1):73–86, 1994.

[28] J. Krajíček. On the weak pigeonhole principle. *Fundamenta Mathematicae* 170(1-3):123–140, 2001.

[29] J. Krajíček. Combinatorics of first order structures and propositional proof systems. *Archive for Mathematical Logic* 43(4):427–441, 2004.

[30] S. Kreutzer and S. Ordyniak. Digraph decompositions and monotonicity in digraph searching. *Theoretical Computer Science* 412(35):4688–4703, 2011.

[31] Y. Li, A. A. Razborov and B. Rossman. On the AC0 Complexity of Subgraph Isomorphism. 55th IEEE Symposium on Foundations of Computer Science, pages 344–353, 2014.

[32] A. Maciel, T. Pitassi, and A. R. Woods. A new proof of the weak pigeonhole principle. *Journal of Computer and System Sciences* 64(4):843–872, 2002.

[33] D. Marx. Can you beat treewidth? *Theory of Computing* 6(1):85–112, 2010.

[34] R. Mazala. Infinite games. In E. Grädel and W. Thomas, editors, *Automata, Logics, and Infinite Games*, Lecture Notes in Computer Science 2500, chapter 2. Springer, 2002.

[35] M. Müller and S. Szeider. Revisiting space in proof complexity: treewidth and pathwidth. 38th Mathematical Foundations of Computer Science, Lecture Notes in Computer Science 8087, pages 704–716, Springer, 2013.

[36] J. Nordström. Narrow proofs may be spacious: separating space and width in resolution. *SIAM Journal on Computing* 39(1):59–121, 2009.

[37] J. Nordström. On the interplay between proof complexity and SAT solving. *ACM SIGLOG News* 2(3):19–44, 2015.

[38] J. Paris and A. Wilkie. Counting problems in bounded arithmetic. *Methods in mathematical logic (Caracas, 1983)*, Lecture Notes in Mathematics 1130, pages 317–340. Springer, 1985.

[39] A. A. Razborov. Proof complexity of pigeonhole principles. 5th Developments in Language Theory, Lecture Notes in Computer Science 2295, pages 100–116, 2002.

[40] A. A. Razborov. Resolution lower bounds for the weak functional pigeonhole principle. *Theoretical Computer Science* 303(1):233–243, 2003.

[41] A. A. Razborov. Pseudorandom generators hard for *k*-DNF resolution and polynomial calculus. *Annals of Mathematics* 181:415–472, 2015.

[42] A. A. Razborov. A new kind of trade-offs in propositional proof complexity. ECCC TR15-033, 2015. *Journal of the ACM*, to appear.

[43] S. Riis. A complexity gap for tree resolution. *Computational Complexity* 10(3):179–209, 2001.

[44] N. Segerlind. Exponential separation between Res(k) and Res($k+1$) for $k \leq \varepsilon \log n$. *Information Processing Letters* 93(4): 185–190. 2005.
[45] N. Segerlind. The complexity of propositional proofs. *Bulletin of Symbolic Logic* 13(4):417–481, 2007.
[46] N. Segerlind, S. Buss, and R. Impagliazzo. A switching lemma for small restrictions and lower bounds for k-DNF resolution. *SIAM Journal on Computing* 33(5):1171–1200, 2004.
[47] G. Stålmarck. Short resolution proofs for a sequence of tricky formulas. *Acta Informatica* 33(3):277–280, 1996.