On Digraph Coloring Problems and Treewidth Duality

Albert Atserias*
Universitat Politècnica de Catalunya
Barcelona, Spain
atserias@lsi.upc.edu

November 24, 2006

Abstract

It is known that every constraint-satisfaction problem (CSP) reduces, and is in fact polynomially equivalent, to a digraph coloring problem. By carefully analyzing the constructions, we observe that the reduction is quantifier-free. Using this, we illustrate the power of the logical approach to CSPs by resolving two conjectures about treewidth duality in the digraph case. The point is that the analogues of these conjectures for general CSPs were resolved long ago by proof techniques that break down for digraphs. We also completely characterize those CSPs that are first-order definable and show that they coincide with those that have finitary tree duality. The combination of this result with an older result by Nešetřil and Tardif shows that there is a computable listing of all template structures whose CSP is definable in full first-order logic. Finally, we provide new width lower bounds for some tractable CSPs. The novelty is that our bounds are a tight function of the treewidth of the underlying instance. As a corollary we get a new proof that there exist tractable CSPs without bounded treewidth duality.

*This work was partially supported by CICYT TIN2004-04343 and by the European Commision through the RTN COMBSTRU HPRN-CT2002-00278.

1 Introduction

Let $\mathbf{H}$ be a graph, or more generally, a finite relational structure. Consider the following computational problem called $\mathbf{H}$-coloring: "Given a finite structure $\mathbf{G}$ of the same type as $\mathbf{H}$, is there a homomorphism from $\mathbf{G}$ to $\mathbf{H}$?". Recall that a homomorphism is a mapping from the elements of $\mathbf{G}$ to the elements of $\mathbf{H}$ that preserves the relations. If such a mapping exists we say that $\mathbf{G}$ is $\mathbf{H}$-colorable. The problem is so called because if $\mathbf{H}$ is the complete graph with k vertices $\mathbf{K}_k$, then a graph $\mathbf{G}$ is $\mathbf{H}$-colorable if and only if it is k-colorable in the standard sense of graph theory. For a general structure $\mathbf{B}$, the $\mathbf{B}$-coloring problem is also called $\text{CSP}(\mathbf{B})$, for *constraint-satisfaction problem*, since the relations in $\mathbf{B}$ can now encode arbitrary constraints. Thus, $\text{CSP}(\mathbf{B})$ problems greatly generalize graph coloring problems, and are thus of major relevance for graph theory and some areas of computer science.

A good question at this point is this: for what $\mathbf{B}$'s is $\text{CSP}(\mathbf{B})$ tractable? As it turns out, for graphs this question is completely resolved by the following beautiful Theorem:

Theorem 1 ([11]) *Let $\mathbf{H}$ be a graph. Then,*

1. *if $\mathbf{H}$ is bipartite, then $\text{CSP}(\mathbf{H})$ is in P,*
2. *if $\mathbf{H}$ is non-bipartite, then $\text{CSP}(\mathbf{H})$ is NP-complete.*

Thus, the complexity of the CSP is completely classified in the case of graphs. Unfortunately, no such classification result is known for arbitrary structures, not even for directed graphs, and in fact the classification problem appears to be a difficult one. The so-called Dichotomy Conjecture of Feder and Vardi has led the recent research in this important area. The conjecture aims for a complete classification as in the Theorem of Hell and Nešetřil above:

Conjecture 1 ([9]) *Let $\mathbf{B}$ be a finite relational structure. Then,*

1. *either $\text{CSP}(\mathbf{B})$ is in P,*
2. *or $\text{CSP}(\mathbf{B})$ is NP-complete.*

Let us note that the special case of the conjecture for directed graphs (digraphs) remains open as well. As a matter of fact, Feder and Vardi proved that there is a dichotomy for digraphs if and only if there is one for arbitrary structures. This latter result, among other reasons, will motivate our focussing on digraphs in certain parts of this paper.

The importance of the Dichotomy Conjecture may not be so much what it literally says, namely that all CSPs are in P or NP-complete and that there is nothing in between, but the likely event that a satisfactory proof would provide deep algorithmic understanding of all tractable CSPs. Conversely, a satisfactory counterexample to the conjecture (modulo P $\neq$ NP or a related assumption) would populate the rather meager class of problems of *intermediate complexity*.

1.1 The unification project

What makes a CSP($\mathbf{B}$) tractable? The first attempt towards a unifying explanation is perhaps due to Gutjahr, Welzl, and Woeginger [10] who showed that if a digraph CSP enjoys the so-called *X-underbar property*, then it is tractable via the so-called arc-consistency algorithm. Subsequently, many known cases in the literature were shown to enjoy the X-underbar property. It was later noticed by Hell, Nešetřil and Zhu [14] that every CSP($\mathbf{B}$) having the X-underbar property also has *bounded treewidth duality*, and that the problems of the second sort are tractable via the bounded-width consistency algorithm. In a nutshell, CSP($\mathbf{B}$) enjoys bounded treewidth duality if there exists a collection of structures $\mathcal{T}$ of bounded treewidth that *obstruct* in the following sense:

$$\mathbf{A} \not\to \mathbf{B} \Longleftrightarrow \mathbf{T} \to \mathbf{A} \text{ for some } \mathbf{T} \in \mathcal{T}.$$

Here we use the notation $\mathbf{A} \to \mathbf{B}$ to denote the existence of a homomorphism from $\mathbf{A}$ to $\mathbf{B}$, and $\mathbf{A} \not\to \mathbf{B}$ to denote non-existence.

But the systematic study and classification of tractable cases of general CSPs started with the influential work of Feder and Vardi [9]. A key observation made in [9] is that many tractable cases are explained by the definability of CSP($\mathbf{B}$) in Datalog, and as it turns out, that expressibility in Datalog, solvability by the bounded-width consistency algorithm, and enjoying bounded treewidth duality are equivalent concepts. These are important consequences of the logical approach to CSPs pionnered in [9] and further developped in [17]. A different unifying framework also introduced in [9] comes from group theory, and subsequent work has shown that this is better explained by universal algebra methods [15, 4]. In brief, the current state of affairs in the unification project is the following: every CSP that is known to be tractable falls into one of two different categories: either (a) it enjoys bounded treewidth duality and is solvable by the bounded-width consistency algorithm, which is the same, or (b) there is a polynomial-time algorithm by universal algebra methods that generalize Gaussian elimination for systems of linear equations. That all CSPs fall in one of these two cases is essentially the thesis raised in [9], with case (b) evolving from group theory to universal algebra.

It is important to note that, prior to the present work, case (b) above appeared to differ from (a) only for vocabularies having at least one relation symbol of arity at least three. Indeed, Feder and Vardi [9] showed that there are CSP problems that are tractable due to (b) but not to (a) by showing that the bounded-width consistency algorithm fails on them. For their proof technique, however, it was essential that the structure incorporates a ternary relation with the so-called "ability to count". But are there tractable cases other than (a) in the absence of ternary relations?

The common theme of this article is the observation that the logical approach to CSPs is a rather powerful tool to obtain solutions to these and other related questions from the literature. In fact, all our results show how the logical approach provides a convenient language to combine many previously known arguments to get new results. We use three main technical tools in our development: (1) the concept of quantifier-free reduction, to study the reduction to the digraph case, (2) Gaifman's locality theorem for first-order logic in combination with hypergraphs of large girth and large chromatic number, which is useful to get preservation results, and (3) inexpressibility results through pebble game arguments and treewidth, which is useful to analyze the power of consistency algorithms. We introduce and illustrate each of these tools through three concrete applications.

1.2 Application: digraphs vs. general structures

From Theorem 1, on graphs, every tractable $\mathrm{CSP}(\mathbf{B})$ is solvable by the bounded-width algorithm, unless P = NP. This is because when $\mathbf{B}$ is a bipartite graph, the bounded-width algorithm works for $\mathrm{CSP}(\mathbf{B})$. Moreover, Nešetřil and Zhu [23] succeeded in showing, without any complexity assumption, the converse to this: if $\mathbf{B}$ is not bipartite, then the bounded-width algorithm does not work for $\mathrm{CSP}(\mathbf{B})$. This is quite satisfactory. For digraphs, however, the situation is dramatically different. While prior to the present work no single $\mathrm{CSP}(\mathbf{B})$ for a digraph $\mathbf{B}$ appeared to be tractable for a reason other than bounded-width, no proof of this was known even assuming P $\neq$ NP. This annoying situation is the consequence of a missing classification theory for digraphs, and motivated Hell, Nešetřil and Zhu to pose the following question:

> **Question [14]:** *Are there tractable* $\mathrm{CSP}(\mathbf{B})$*, for a digraph* $\mathbf{B}$*, that do not enjoy bounded treewidth duality?*

The point of this question is that its analogue for general CSPs was resolved long ago by proof techniques that break down for digraphs. These are the results about the "ability to count" in [9] mentioned above. In the first part of this paper we answer precisely this question, in the affirmative, without any complexity assumption. The resulting digraph has exactly 368 vertices and 432 edges. Our proof is not difficult but is interesting as it proves the power of the logical approach. It consists in verifying that the known reductions from general structures to digraphs can be made quantifier-free. Of course, details are required (how do we get rid of negations?) and will be provided in Section 4.

Following along these lines, we are also able to show that there is a *triad* $\mathbf{T}$, a special kind of oriented tree with a single node of degree three, for which CSP$(\mathbf{T})$ provably does not have bounded treewidth duality. This is again the first such example for oriented trees, without assuming P $\neq$ NP, and resolves another question raised by Hell, Nešetřil and Zhu. Let us point out that, in recent work, Larose and Zádori [19] also find certain interesting digraphs having unbounded treewidth duality. The techniques are different, but the questions are similar.

1.3 Application: finite dualities

Since the Dichotomy Conjecture appears to be a difficult question, it is perhaps a good idea to consider special cases. Consider, for example, the following question raised by Feder and Vardi:

> **Question [9]**: *Which* CSP$(\mathbf{B})$ *are definable in Datalog?*

We mentioned already that these are exactly those enjoying bounded treewidth duality. But is it decidable whether a problem has bounded treewidth duality? It turns out that this question is open except for the special case of trees (treewidth one) [12]. A different but related question was considered by Nešetřil and Tardif [22] (see also [20, Theorem 3.13]); they characterized those CSP$(\mathbf{B})$ having duality with a finite obstruction set $\mathcal{T}$. In turn, it is not hard to see that these are precisely the CSP$(\mathbf{B})$ problems that are definable in the recursion-free fragment of Datalog, or equivalently, in the existential-positive fragment of first-order logic. A natural question arises at this point (first raised in [7]):

> **Question [7]:** *Which* CSP$(\mathbf{B})$ *are definable in full first-order logic?*

In Section 5 we answer this question. Using the methods in [3], we show that CSP$(\mathbf{B})$ is definable in first-order logic if and only if it is definable in the existential-positive fragment of first-order logic. In turn, together with the results by Nešetřil and Tardif above, this provides a semi-decidable classification (note that the obvious classification is only in Σ_2, the second level of the arithmetic hierarchy). We refer the reader to [7] for similar results concerning first-order definability in homomorphism problems, and to [18] for very recent results that give a decidable characterization of problems with finitary duality. Building on our result, this fully settles the issue of characterizing those CSP($\mathbf{B}$) that are first-order definable.

Let us point out that our result follows from a solution to the well-known Homomorphism-Preservation Conjecture in finite model theory [25], but our proof does not rely on it and is based entirely on locality arguments for first-order logic.

1.4 Application: new width bounds based on treewidth

To complete the picture, in Section 6 we provide a new proof that there are tractable CSPs, with at least one ternary relation symbol, that are not definable in Datalog. This also has the nice feature of making our results about digraph CSPs self-contained.

Our proof is a pebble-game argument for Tseitin systems analogous to those used for proving width lower bounds for random formulas [2]. The novelty in our result is that we obtain width bounds as a tight function of the treewidth of the underlying graph while previous width lower bounds were based on expansion, a more demanding parameter that can be significantly smaller than treewidth. We believe this is the more interesting aspect of this result. Another interesting point is that, as opposed to the proof in [9], our proof does not use the lower bounds for monotone circuits for perfect matching due to Razborov. Instead, we use the Robber-Cops game characterization of treewidth due to Seymour and Thomas [27].

2 Preliminaries

2.1 Structures, homomorphisms, graphs, and hypergraphs

A finite relational vocabulary σ is a finite collection of relation symbols each of a specified positive arity. A *σ-structure* $\mathbf{A}$ is a *universe* A together with an *interpretation* $R^{\mathbf{A}} \subseteq A^r$ for every relation symbol $R \in \sigma$ of arity r. We write $\mathbf{A} = (A, (R^{\mathbf{A}})_{R\in\sigma})$. All our vocabularies and structures in this paper are FINITE. Thus, we often omit saying it.

If $\mathbf{A}$ and $\mathbf{B}$ are σ-structures, a *homomorphism* is a mapping $h : A \to B$ such that if $(a_1, \ldots, a_r) \in R^{\mathbf{A}}$, then $(h(a_1), \ldots, h(a_r)) \in R^{\mathbf{B}}$. We write $\mathbf{A} \to \mathbf{B}$ to denote the existence of a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. We say that $\mathbf{A}$ is a *substructure of* $\mathbf{B}$, denoted by $\mathbf{A} \subseteq \mathbf{B}$, if $A \subseteq B$ and the inclusion mapping is a homomorphism. We say that $\mathbf{A}$ is an *induced substructure of* $\mathbf{B}$ if $R^{\mathbf{B}} \cap A^r = R^{\mathbf{A}}$ for every $R \in \sigma$ of arity r. If $\mathbf{A} \subseteq \mathbf{B}$ and $\mathbf{A} \neq \mathbf{B}$ we say that $\mathbf{A}$ is a *proper substructure of* $\mathbf{B}$. The *direct product* of two σ-structures $\mathbf{A}$ and $\mathbf{B}$ has the cartesian product $A \times B$ as universe, and the relation R is interpreted as the set of tuples $((a_1, b_1), \ldots, (a_r, b_r))$ such that $(a_1, \ldots, a_r) \in R^{\mathbf{A}}$ and $(b_1, \ldots, b_r) \in R^{\mathbf{B}}$.

A *hypergraph* $\mathbf{H} = (V, E)$ is a set of vertices V together with a set E of subsets of V called hyperedges, or simply edges. An edge is called a loop if it is a singleton. A walk in a hypergraph is an alternating sequence of vertices and edges

$$u_0, e_0, u_1, \ldots, u_m, e_m, u_{m+1}$$

such that for every $i \in \{0, \ldots, m\}$ we have that $u_i \neq u_{i+1}$ and $\{u_i, u_{i+1}\} \subseteq e_i$. The walk is a path if $e_i \neq e_j$ and $u_i \neq u_j$ whenever $0 \leq i < j \leq m$, and a cycle if, in addition, $u_0 = u_{m+1}$. This sort of cycle in a hypergraph is sometimes called a Berge-cycle. The length of a walk, path, or cycle as above is m, the number of traversed edges. The girth of a hypergraph is the length of the shortest cycle. Note that the girth is always at least two. If the hypergraph has no cycles, we call it acyclic and its girth is infinite.

A *graph* $\mathbf{G} = (V, E)$ is hypergraph where every edge has exactly two vertices. We may view graphs as structures for the vocabulary of a single binary relation symbol E whose interpretation is irreflexive and symmetric. In this case we sometimes write $\{u, v\} \in E^{\mathbf{G}}$ instead of $(u, v) \in E^{\mathbf{G}}$. A *digraph* is just a structure for the vocabulary $\sigma = (E)$, where E is a binary relation symbol. An oriented graph is a digraph such that $E^{\mathbf{G}}$ is irreflexive and anti-symmetric.

2.2 First-order logic, Datalog, and infinitary logic

Let σ be a relational vocabulary. Let $x_1, x_2, \ldots$ be *first-order variables*. The *atomic formulas* are of the form $R(y_1, \ldots, y_r)$ or $y_1 = y_2$, where $R \in \sigma$ is a relation symbol of arity r and each y_i is a first-order variable. The collection of *first-order formulas*, denoted by FO, is obtained by closing the atomic formulas under negation, conjunction, disjunction, and existential and universal quantification. The existential-positive fragment of FO, denoted by $\exists$FO, is obtained by closing the atomic formulas under conjunction, disjunction, and existential quantification. The conjunctive existential-positive fragment of FO, denoted $\wedge\exists$FO, is obtained by closing the atomic formulas under conjunction and existential quantification.

Although we mention Datalog in this paper, we will not really work with it, so we do not define it. Instead we will need infinitary logic. The collection of *infinitary formulas*, denoted by $\mathrm{L}_{\infty\omega}$, is obtained by closing the atomic formulas under negation, infinitary disjunction, infinitary conjunction, and existential and universal quantification. The existential-positive fragment of $\mathrm{L}_{\infty\omega}$, denoted by $\exists\mathrm{L}_{\infty\omega}$, is obtained by closing the atomic formulas under infinitary disjunction, infinitary conjunction, and existential quantification.

If L is a collection of formulas, we write L^k for the collection of such formulas that use only the variables $x_1, \ldots, x_k$. Thus $\exists\mathrm{FO}^k$ is the k-variable fragment of

$\exists$FO and $\exists L^k_{\infty\omega}$ is the k-variable fragment of $\exists L_{\infty\omega}$. We write $\neg L$ for the collection of formulas of the form $\neg\varphi$, with $\varphi \in L$. We write $L(\neg, \neq)$ when, in addition to all atomic formulas, we also allow negated atomic formulas and inequalities in the formation rules. Thus, $\neg\exists L^k_{\infty\omega}(\neg, \neq)$ denotes the negations of formulas in the k-variable fragment of the closure of the atomic and negated atomic formulas under infinitary disjunction, infinitary conjunction, and existential quantification.

Let us note here that existential-positive formulas are closely related to homomorphisms. Indeed, if φ is an existential-positive formula that holds in $\mathbf{A}$ and $\mathbf{A} \to \mathbf{B}$, then φ also holds in $\mathbf{B}$. Moreover, for every finite σ-structure $\mathbf{C}$, there exists a $\wedge\exists$FO-sentence $\varphi_{\mathbf{C}}$, called the *canonical query of* $\mathbf{C}$, such that for every structure $\mathbf{A}$ we have that

$$\mathbf{A} \models \varphi_{\mathbf{C}} \iff \mathbf{C} \to \mathbf{A}.$$

The formula $\varphi_{\mathbf{C}}$ is obtained by viewing each point of $\mathbf{C}$ as a first-order variable, viewing each tuple $(c_1, \ldots, c_r) \in R^{\mathbf{C}}$ as an atomic formula $R(c_1, \ldots, c_r)$. The formula is formed by conjoining all those atomic formulas together and then existentially quantifying all variables.

2.3 Pebble games

The existential k-pebble game is played by two players, the Spoiler and the Duplicator, on a board that consists of two σ-structures $\mathbf{A}$ and $\mathbf{B}$. The goal of the Spoiler is to prove that $\mathbf{A} \not\to \mathbf{B}$, while the goal of the Duplicator is to convince the Spoiler that $\mathbf{A} \to \mathbf{B}$. Each player has k pebbles numbered $1, \ldots, k$ that, at the begining of each round in the game, each may either be unused or rest on a point of $\mathbf{A}$ or $\mathbf{B}$. Initially, all pebbles are unused. At each round of the game, the Spoiler either places an unused pebble i over a point a_i of $\mathbf{A}$, or removes a used pebble i from the point it rests on $\mathbf{A}$. Then the Duplicator must respond by either placing her corresponding pebble i over a point b_i of $\mathbf{B}$, or removing her corresponding pebble i from the point it rests on $\mathbf{B}$. If at the end of the round the mapping that sends a_i to b_i, for every used pebble i, is not a partial homomorphism from $\mathbf{A}$ and $\mathbf{B}$, the Spoiler wins the game. Otherwise the game proceeds to the next round. If the Duplicator has a strategy to play forever, we say that the Duplicator wins.

The definition can be made formal through systems of partial homomorphisms. Recall that a partial map $h : A \to B$ is a partial homomorphism from $\mathbf{A}$ to $\mathbf{B}$ if it is a homomorphism from the substructure of $\mathbf{A}$ induced by the domain $\mathrm{Dom}(h)$ of h to $\mathbf{B}$. We say that h is a partial k-homomorphism if the size of $\mathrm{Dom}(h)$ is at most k.

Definition 1 ([16, 17]) *Let $\mathbf{A}$ and $\mathbf{B}$ be two σ-structures and let $k \geq 1$ be an integer. A winning strategy for the Duplicator in the existential k-pebble game on*

$\mathbf{A}$ *and* $\mathbf{B}$ *is a non-empty collection* $\mathcal{F}$ *of partial* k*-homomorphisms from* $\mathbf{A}$ *to* $\mathbf{B}$ *satisfying the following properties:*

1. *(subfunction property) if* $g \in \mathcal{F}$ *and* $f \subseteq g$*, then* $f \in \mathcal{F}$*,*
2. *(forth property) if* $g \in \mathcal{F}$ *with* $|\mathrm{Dom}(g)| < k$ *and* $a \in A - \mathrm{Dom}(g)$*, then there exists* $h \in \mathcal{F}$ *such that* $g \subseteq h$ *and* $a \in \mathrm{Dom}(h)$*.*

If there exists a winning strategy for the Duplicator we say that the Duplicator wins the existential k*-pebble game on* $\mathbf{A}$ *and* $\mathbf{B}$ *and we write* $\mathbf{A} \leq^k_{\infty\omega} \mathbf{B}$*.*

The main property of the existential k-pebble game is that it characterizes definability in the logic $\exists\mathrm{L}^k_{\infty\omega}$ as proved by Kolaitis and Vardi [16, 17]. For the following Theorem, the fact that $\mathbf{A}$ and $\mathbf{B}$ are finite is essential to get equivalences with (3) and (4):

Theorem 2 ([16, 17]) *Let* σ *be a relational vocabulary of maximum arity* r *and let* $k \geq r$ *be an integer. Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be finite* σ*-structures. The following are equivalent:*

1. $\mathbf{A} \leq^k_{\infty\omega} \mathbf{B}$*,*
2. *every* $\exists\mathrm{L}^k_{\infty\omega}$ *sentence that holds in* $\mathbf{A}$ *also holds in* $\mathbf{B}$*,*
3. *every* $\exists\mathrm{FO}^k$ *sentence that holds in* $\mathbf{A}$ *also holds in* $\mathbf{B}$*,*
4. *every* $\wedge\exists\mathrm{FO}^k$ *sentence that holds in* $\mathbf{A}$ *also holds in* $\mathbf{B}$*.*

The existential k-pebble game can be modified to characterize the expressive power of $\exists\mathrm{L}^k_{\infty\omega}(\neg, \neq)$. The idea is to require the positions of the game to be partial isomorphisms instead of partial homomorphisms. The rest of the game is the same. We write $\mathbf{A} \leq^{k,\mathrm{iso}}_{\infty\omega} \mathbf{B}$ if the Duplicator wins the modified game on $\mathbf{A}$ and $\mathbf{B}$. We get the analogue of Theorem 2 when we replace all fragments L by $L(\neg, \neq)$ and $\leq^k_{\infty\omega}$ by $\leq^{k,\mathrm{iso}}_{\infty\omega}$.

Theorem 3 *Let* σ *be a relational vocabulary of maximum arity* r *and let* $k \geq r$ *be an integer. Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be finite* σ*-structures. The following are equivalent:*

1. $\mathbf{A} \leq^{k,\mathrm{iso}}_{\infty\omega} \mathbf{B}$*,*
2. *every* $\exists\mathrm{L}^k_{\infty\omega}(\neg, \neq)$ *sentence that holds in* $\mathbf{A}$ *also holds in* $\mathbf{B}$*,*
3. *every* $\exists\mathrm{FO}^k(\neg, \neq)$ *sentence that holds in* $\mathbf{A}$ *also holds in* $\mathbf{B}$*,*
4. *every* $\wedge\exists\mathrm{FO}^k(\neg, \neq)$ *sentence that holds in* $\mathbf{A}$ *also holds in* $\mathbf{B}$*.*

2.4 Treewidth

Let $\mathbf{G} = (V, E)$ be a graph. A *tree* is a connected acyclic graph. A *forest* is an acyclic graph. A *tree-decomposition* of $\mathbf{G}$ is a labeled tree $\mathbf{T}$ such that (1) every node of $\mathbf{T}$ is labeled by a non-empty subset of V, (2) for every edge $\{u, v\} \in E$ there is a node of $\mathbf{T}$ whose label contains $\{u, v\}$, and (3) for every $u \in V$, the set X of nodes of $\mathbf{T}$ whose labels include u is a connected subtree of $\mathbf{T}$. The *width* of a tree-decomposition is the maximum cardinality of a label in $\mathbf{T}$ minus one. The *treewidth* of $\mathbf{G}$ is the smallest k for which $\mathbf{G}$ has a tree-decomposition of width k. The *treewidth* of a σ-structure is the treewidth of its Gaifman graph, where the *Gaifman graph*, denoted by $\mathcal{G}(\mathbf{A})$, is the graph with vertices A and with an edge between distinct a and b if they appear together in some tuple of $\mathbf{A}$. The collection of σ-structures of treewidth at most k is denoted by $\mathcal{T}_k(\sigma)$ or $\mathcal{T}_k$ if σ is understood from context.

There is a tight and surprising connection between treewidth and k-variable logics as made clear by the following result. Recall that $\varphi_{\mathbf{C}}$ denotes the canonical conjunctive query of $\mathbf{C}$ and has the property that, for every structure $\mathbf{A}$, we have $\mathbf{A} \models \varphi_{\mathbf{C}}$ if and only if $\mathbf{C} \to \mathbf{A}$.

Theorem 4 ([17, 6]) *Let σ be a relational vocabulary of maximum arity r and let $k \geq r$ be an integer. The following two implications hold:*

1. *for every $\mathbf{C} \in \mathcal{T}_k(\sigma)$, there exists $\psi \in \wedge\exists\mathrm{FO}^{k+1}$ such that $\psi \equiv \varphi_{\mathbf{C}}$,*
2. *for every $\psi \in \wedge\exists\mathrm{FO}^{k+1}$, there exists $\mathbf{C} \in \mathcal{T}_k(\sigma)$ such that $\varphi_{\mathbf{C}} \equiv \psi$.*

Roughly speaking, the idea behind of the proof of this result is that parse-trees of $\wedge\exists\mathrm{FO}^{k+1}$-sentences and tree-decompositions of width k are more or less interchangeable concepts when we view variables as points of a structure, or points of a structure as variables. For one thing, the subformulas of such sentences always have at most $k+1$ free variables because there are at most $k+1$ variables in total. Thus, the sets of free variables of the subformulas of the parse-tree play the role of bags in the tree-decomposition. Note that the connectivity condition of free variables is satisfied trivially in parse-trees.

3 Duality and Quantifier-free Reductions

Let σ be a relational vocabulary. Let $\mathbf{B}$ be a σ-structure and let $\mathcal{D}$ be a collection of σ-structures. We say that $\mathcal{D}$ is a duality for $\mathrm{CSP}(\mathbf{B})$ if for every σ-structure $\mathbf{A}$ we have

$$\mathbf{A} \not\to \mathbf{B} \iff \mathbf{C} \to \mathbf{A} \text{ for some } \mathbf{C} \in \mathcal{D}.$$

If CSP$(\mathbf{B})$ has a duality $\mathcal{D}$ that is finite, we say that CSP$(\mathbf{B})$ has finitary duality. When $\mathcal{D} \subseteq \mathcal{T}_k(\sigma)$, we say that CSP$(\mathbf{B})$ has treewidth-k duality.

3.1 Bounded width and logic

The following result is the culmination of a successful line of research in CSPs by logical methods.

Theorem 5 ([9, 17, 8]) *Let σ be a relational vocabulary of maximum arity r, let $k \geq r$ be an integer, and let $\mathbf{B}$ be a σ-structure. Then, the following are equivalent:*

1. CSP$(\mathbf{B})$ *has treewidth-k duality,*
2. CSP$(\mathbf{B})$ *is definable in* $\neg(k+1)$-DATALOG*,*
3. CSP$(\mathbf{B})$ *is definable in* $\neg(k+1)$-DATALOG$(\neg, \neq)$*,*
4. CSP$(\mathbf{B})$ *is definable in* $\neg\exists\mathrm{L}^{k+1}_{\infty\omega}$*,*
5. CSP$(\mathbf{B})$ *is definable in* $\neg\exists\mathrm{L}^{k+1}_{\infty\omega}(\neg, \neq)$.
6. CSP$(\mathbf{B}) = \{\mathbf{A} : \mathbf{A} \leq^{k+1}_{\infty\omega} \mathbf{B}\}$.

The equivalence between (1) and (2) was sketched in [9, Theorem 23]. We do not really need it anyway. It is obvious that (2) implies (3) and that (4) implies (5). The equivalences between (2), (4) and (6) are from [17, Theorem 4.8]. The implication (5) to (4) follows from [8, Theorem 1], which is a recent result that we also prove below for completeness. Hence, when we prove below that (1) and (6) are equivalent we are showing that (1), (2), (4), (5), and (6) are equivalent. It would be left to prove that (3) is equivalent to any of the others, say by showing that it implies (5). But since this is easily seen from the techniques in [17] and we do not need it for the rest of the paper, we will omit the proof. At any rate, the statements and proofs of the equivalences expressed in Theorem 5 are spread all over the literature and we think it is useful to spell them out explicitly here.

Proof of equivalence between (1) and (6) in Theorem 5: Suppose (1) holds. That is, CSP$(\mathbf{B})$ has treewidth-k duality witnessed by $\mathcal{D} \subseteq \mathcal{T}_k$. We show (6). Clearly, if $\mathbf{A} \to \mathbf{B}$ then $\mathbf{A} \leq^{k+1}_{\infty\omega} \mathbf{B}$. Suppose now that $\mathbf{A} \not\to \mathbf{B}$. Then there exists $\mathbf{C} \in \mathcal{D}$ such that $\mathbf{C} \to \mathbf{A}$. Of course $\mathbf{C} \not\to \mathbf{B}$ as well. Apply Theorem 4 to $\mathbf{C}$ and get a $\wedge\exists\mathrm{FO}^{k+1}$-sentence ψ that holds in $\mathbf{A}$ and not in $\mathbf{B}$. It follows from Theorem 2 that $\mathbf{A} \not\leq^{k+1}_{\infty\omega} \mathbf{B}$.

Suppose (6) holds. Let $\mathcal{D}$ be the collection of all σ-structures of treewidth at most k that do not map to $\mathbf{B}$. We claim that $\mathcal{D}$ is a duality for CSP$(\mathbf{B})$. Suppose first that $\mathbf{A} \not\to \mathbf{B}$. Then $\mathbf{A} \not\leq^{k+1}_{\infty\omega} \mathbf{B}$, so by Theorem 2 there exists a $\wedge\exists\mathrm{FO}^{k+1}$-sentence

ψ that holds in $\mathbf{A}$ and not in $\mathbf{B}$. Apply Theorem 4 to ψ and get a σ-structure $\mathbf{C}$ of treewidth at most k such that $\mathbf{C} \to \mathbf{A}$ and $\mathbf{C} \not\to \mathbf{B}$. This means that $\mathbf{C} \to \mathbf{A}$ for some $\mathbf{C} \in \mathcal{D}$ and the proof is complete. □

We still need to prove the implication (5) to (4). This is a key fact that follows from a recent result by Feder and Vardi [8]. Although the proof is the same as in [8], we include here, for completeness, a slightly stronger statement with its proof:

Theorem 6 ([8]) *Let σ be a relational vocabulary of maximum arity r, let $k \geq r$ be an integer, let $\mathcal{C}$ be a class of σ-structures that is closed under direct products, and let $Q \subseteq \mathcal{C}$ be preserved under homomorphisms on $\mathcal{C}$. Then the following are equivalent:*

1. *Q is definable in $\exists\mathrm{L}^k_{\infty\omega}$ on $\mathcal{C}$,*
2. *Q is definable in $\exists\mathrm{L}^k_{\infty\omega}(\neq, \neg)$ on $\mathcal{C}$.*

Proof: Obviously (1) implies (2). We show (2) implies (1). Suppose Q is not definable in $\exists\mathrm{L}^k_{\infty\omega}$ on $\mathcal{C}$. Then, by Theorem 2, there exists $\mathbf{A} \in \mathcal{C}$ and $\mathbf{B} \in \mathcal{C}$ such that $\mathbf{A} \in Q$ and $\mathbf{B} \notin Q$, yet $\mathbf{A} \leq^k_{\infty\omega} \mathbf{B}$. Otherwise the $\exists\mathrm{L}^k_{\infty\omega}$-sentence

$$\bigvee_{\mathbf{A}} \bigwedge_{\varphi} \varphi$$

would define Q on $\mathcal{C}$, where $\mathbf{A}$ ranges over Q in the disjunction, and φ ranges over $\{\varphi \in \exists\mathrm{L}^k_{\infty\omega} : \mathbf{A} \models \varphi\}$ in the conjunction. We will show that $\mathbf{A} \leq^{k,\mathrm{iso}}_{\infty\omega} \mathbf{A} \times \mathbf{B}$. Note that $\mathbf{A} \times \mathbf{B}$ belongs to $\mathcal{C}$ but not to Q. Indeed, $\mathbf{A} \times \mathbf{B} \to \mathbf{B}$ and Q is preserved under homomorphisms on $\mathcal{C}$, so $\mathbf{A} \times \mathbf{B} \notin Q$ follows from $\mathbf{B} \notin Q$. The result will follow from Theorem 3.

We are assuming that $\mathbf{A} \leq^k_{\infty\omega} \mathbf{B}$. Let $\mathcal{F}$ be a winning strategy for the Duplicator. For every $f \in \mathcal{F}$, define a partial map $f' : A \to A \times B$ as $f'(a) = (a, f(a))$ for every $a \in \mathrm{Dom}(f)$. Let $\mathcal{F}' = \{f' : f \in \mathcal{F}\}$. It is straightforward to check that $\mathcal{F}'$ witnesses that $\mathbf{A} \leq^{k,\mathrm{iso}}_{\infty\omega} \mathbf{A} \times \mathbf{B}$. □

This completes the proof of Theorem 5. In case any of the equivalent conditions of the Theorem holds, and so all hold, we say that CSP($\mathbf{B}$) has width at most k.

Using the key fact that negations in the atomic formulas do not add power as we just showed, let us now note that bounded width is preserved downward through quantifier-free reductions. First we need to define this concept.

3.2 Quantifier-free reductions and preservation

Let σ and $\tau = (R_1, \ldots, R_s)$ be two relational vocabularies. A *k-ary first-order interpretation with p parameters of τ in σ* is an $(s+1)$-tuple $\mathbf{I} = (\varphi_U, \varphi_1, \ldots, \varphi_s)$ of first-order formulas over the vocabulary σ, where $\varphi_U = \varphi_U(\mathbf{x}, \mathbf{y})$ has $k+p$ free variables $\mathbf{x} = (x^1, \ldots, x^k)$ and $\mathbf{y} = (y_1, \ldots, y_p)$, and $\varphi_i = \varphi_i(\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{y})$ has $kr + p$ free variables where r is the arity of R_i and each $\mathbf{x}_j = (x_j^1, \ldots, x_j^k)$ and $\mathbf{y} = (y_1, \ldots, y_p)$.

If $\mathbf{A}$ is a σ-structure and $\mathbf{c} = (a_1, \ldots, a_p)$ is a tuple of pairwise different points of $\mathbf{A}$, then the *interpretation of $\mathbf{A}$ through $\mathbf{I}$ with parameters $\mathbf{c}$*, denoted by $\mathbf{I}(\mathbf{A}, \mathbf{a})$, is the τ-structure whose universe is

$$\{\mathbf{a} \in A^k : \mathbf{A} \models \varphi_U(\mathbf{a}, \mathbf{c})\},$$

and whose interpretation for R_i is

$$\{(\mathbf{a}_1, \ldots, \mathbf{a}_r) \in (A^k)^r : \mathbf{A} \models \varphi_U(\mathbf{a}_1, \mathbf{c}) \wedge \cdots \wedge \varphi_U(\mathbf{a}_r, \mathbf{c}) \wedge \varphi_i(\mathbf{a}_1, \ldots, \mathbf{a}_r, \mathbf{c})\}.$$

If each formula in $\mathbf{I}$ is quantifier-free, we say that $\mathbf{I}$ is a *quantifier-free interpretation*.

Now we are ready to define the notion of quantifier-free reduction:

Definition 2 *Let σ and τ be finite relational vocabularies, let $\mathcal{C}$ be a class of σ-structures, and let $\mathcal{D}$ be a class of τ-structures that is closed under isomorphisms. We say that a first-order interpration with p parameters $\mathbf{I}$ of τ in σ is a first-order reduction from $\mathcal{C}$ to $\mathcal{D}$ if for every σ-structure $\mathbf{A}$ with at least p points the following three statements are equivalent:*

1. $\mathbf{A} \in \mathcal{C}$,
2. $\mathbf{I}(\mathbf{A}, \mathbf{c}) \in \mathcal{D}$ *for every proper* $\mathbf{c}$,
3. $\mathbf{I}(\mathbf{A}, \mathbf{c}) \in \mathcal{D}$ *for some proper* $\mathbf{c}$,

where a proper $\mathbf{c}$ is a tuple $(c_1, \ldots, c_p)$ of points of $\mathbf{A}$ such that $c_i \neq c_j$ whenever $i \neq j$. If $\mathbf{I}$ is quantifier-free, then we say that $\mathbf{I}$ is a quantifier-free reduction from $\mathcal{C}$ to $\mathcal{D}$.

In case there is a reduction as in the definition, we say that $\mathcal{C}$ reduces to $\mathcal{D}$. Here is the promised result. The proof is standard but we give it nonetheless.

Lemma 1 *Let σ and σ' be relational vocabularies of arities at most r, and let $k \geq r$ be an integer. Let $\mathbf{B}$ be a σ-structure and let $\mathbf{B}'$ be a σ'-structure. If* CSP$(\mathbf{B})$ *reduces to* CSP$(\mathbf{B}')$ *by a k-ary quantifier-free reduction with p parameters and* CSP$(\mathbf{B}')$ *has width at most k', then* CSP$(\mathbf{B})$ *has width at most $k(k'+1)+p-1$.*

Proof: Let $q = k(k'+1)+p$. We use Theorem 5. Let φ be an $\exists \mathrm{L}^{k'+1}_{\infty\omega}$-sentence over the vocabulary σ' defining the complement of $\mathrm{CSP}(\mathbf{B}')$. We transform every $\exists \mathrm{L}^{k'+1}_{\infty\omega}$-formula ψ into an $\exists \mathrm{L}^{q}_{\infty\omega}(\neg, \neq)$-formula $\mathbf{I}(\psi)$ inductively as follows. First, we make k copies $x^1, \ldots, x^k$ of each variable x of ψ. Let $\mathbf{y} = (y_1, \ldots, y_p)$ be p new variables that will be used as parameters. For the base cases, each atomic formula of the form $R(x_1, \ldots, x_r)$ is transformed to

$$\varphi_R(x_1^1, \ldots, x_1^k, \ldots, x_r^1, \ldots, x_r^k, \mathbf{y}).$$

Each atomic formula of the form $x_1 = x_2$ is transformed to

$$x_1^1 = x_2^1 \wedge \cdots \wedge x_1^k = x_2^k.$$

For the inductive cases, each conjunction $\bigwedge \psi_i$ is transformed to $\bigwedge \mathbf{I}(\psi_i)$. Each disjunction $\bigvee \psi_i$ is transformed to $\bigvee \mathbf{I}(\psi_i)$. And each existentially quantified formula $(\exists x)\psi$ is transformed to

$$(\exists x^1) \cdots (\exists x^k)(\varphi_U(x^1, \ldots, x^k, \mathbf{y}) \wedge \mathbf{I}(\psi)).$$

Now, by induction on the structure of the formula $\psi(x_1, \ldots, x_{k'})$, for every σ-structure $\mathbf{A}$, every proper $\mathbf{c}$, and every $\mathbf{a}_1, \ldots, \mathbf{a}_{k'}$ in the universe of $\mathbf{I}(\mathbf{A}, \mathbf{c})$, we have

$$\mathbf{A} \models \mathbf{I}(\psi)(\mathbf{a}_1, \ldots, \mathbf{a}_{k'}, \mathbf{c}) \Longleftrightarrow \mathbf{I}(\mathbf{A}, \mathbf{c}) \models \psi(\mathbf{a}_1, \ldots, \mathbf{a}_{k'}).$$

In particular, for the sentence φ' we have, for every σ-structure $\mathbf{A}$ and every proper $\mathbf{c}$, that

$$\mathbf{A} \models \mathbf{I}(\varphi')(\mathbf{c}) \Longleftrightarrow \mathbf{I}(\mathbf{A}, \mathbf{c}) \models \varphi'. \tag{1}$$

Finally, we define a sentence φ as follows:

$$(\exists y_1) \cdots (\exists y_p) \left(\bigwedge_{i \neq j} y_i \neq y_j \wedge \mathbf{I}(\varphi') \right).$$

Note that φ is an $\exists \mathrm{L}^{q}_{\infty\omega}(\neg, \neq)$-sentence over the vocabulary σ. Fix now a σ-structure $\mathbf{A}$ with at least p points. We get the following sequence of equivalences:

$$\begin{aligned}
\mathbf{A} \models \neg\varphi \quad &\Longleftrightarrow \quad \mathbf{A} \models \neg \mathbf{I}(\varphi')(\mathbf{c}) \text{ for every proper } \mathbf{c} && (2)\\
&\Longleftrightarrow \quad \mathbf{I}(\mathbf{A}, \mathbf{c}) \models \neg\varphi' \text{ for every proper } \mathbf{c} && (3)\\
&\Longleftrightarrow \quad \mathbf{I}(\mathbf{A}, \mathbf{c}) \to \mathbf{B}' \text{ for every proper } \mathbf{c} && (4)\\
&\Longleftrightarrow \quad \mathbf{A} \to \mathbf{B}. && (5)
\end{aligned}$$

Equivalence (2) follows from the definition of φ. Equivalence (3) follows from (1). Equivalence (4) follows from the fact that $\neg\varphi'$ defines $\mathrm{CSP}(\mathbf{B}')$. And equivalence (5) follows from the fact that $\mathbf{I}$ is a reduction from $\mathrm{CSP}(\mathbf{B})$ to $\mathrm{CSP}(\mathbf{B}')$. We conclude that φ defines the complement of CSP$(\mathbf{B})$ for structures with at least p points. The finitely many exceptions of size at most p can be handled by adding a disjunction of finitely many $\exists\mathrm{FO}^p$-sentences. Hence CSP$(\mathbf{B})$ is definable in $\neg\exists\mathrm{L}^q_{\infty\omega}(\neg,\neq)$, so it has width at most $q-1$ by Theorem 5. □

One expected property of quantifier-free reductions is that they compose. Thus, if $\mathbf{I}$ is a k-ary quantifier-free reduction with p parameters from $\mathcal{C}$ to $\mathcal{D}$ and $\mathbf{J}$ is a k'-ary quantifier-free reduction with p' parameters from $\mathcal{D}$ to $\mathcal{E}$, then there exists a kk'-ary quantifier-free reduction with $kp+p'$ parameters from $\mathcal{C}$ to $\mathcal{E}$.

4 Application: Reduction to digraph CSPs

It is known that every CSP is polynomially equivalent to a digraph coloring problem, which means that for every relational structure $\mathbf{B}$, there exists a digraph $\mathbf{H}$ such that CSP$(\mathbf{B}) \leq^p_m$ CSP$(\mathbf{H})$ and CSP$(\mathbf{H}) \leq^p_m$ CSP$(\mathbf{B})$. The goal of this section is to justify the stronger claim that the reduction from CSP$(\mathbf{B})$ to CSP$(\mathbf{H})$ can be made quantifier-free:

Theorem 7 *Let σ be a relational vocabulary, and let $\mathbf{B}$ be a σ-structure. Then, there exists a digraph $\mathbf{H}$ such that*

1. *there is a quantifier-free reduction from* CSP$(\mathbf{B})$ *to* CSP$(\mathbf{H})$*,*
2. *there is a polynomial-time reduction from* CSP$(\mathbf{H})$ *to* CSP$(\mathbf{B})$*.*

Moreover, $\mathbf{H}$ is a directed acyclic graph of maximum degree 2 and height 3.

Note that the back reduction in (2) need not be quantifier-free. It turns out that in the known proofs of polynomial equivalence between CSP$(\mathbf{B})$ and CSP$(\mathbf{H})$, the reduction from CSP$(\mathbf{B})$ to CSP$(\mathbf{H})$ is already implicitly quantifier-free. Although verifying this is rather tedious and uninteresting, it is necessary to do it at least once, so let us do it nonetheless.

There are two known proofs of equivalence between general CSPs and CSPs with digraph templates. Both, again, come from [9]. The first proof is in Theorems 8 and 10 in [9] and is reconstructued with great care in [12]. It can be checked that (a minor modification of) the reduction from CSP$(\mathbf{B})$ to CSP$(\mathbf{H})$ of that proof is quantifier-free. Here we work out the second known proof. This appears in Theorem 11 in [9] and, to the best of our knowledge, nowhere else. This reduction has the great advantage that, when applied on a specific $\mathbf{B}$, it produces a relatively small and explicit digraph $\mathbf{H}$. For the application we have in mind, our $\mathbf{B}$ will be the template structure of the problem AFFINE-3-SAT which has two elements in its universe and two ternary relations. The resulting digraph $\mathbf{H}$ will have exactly 368 vertices and 432 edges.

4.1 Turning structures into digraphs

Here is the transformation of $\mathbf{B}$ into $\mathbf{H}$. First we transform $\mathbf{B}$ into an intermediate structure $\mathbf{B}'$ of vocabulary $\tau = \{R, S\}$, where R is unary relation simbol, and S is a binary relation symbol. Let $\sigma = (R_1, \dots, R_s)$ be the vocabulary of $\mathbf{B}$. Let $r = r_1 + \cdots + r_s$ be the sum of the arities of the relations in σ. The universe B' of $\mathbf{B}'$ is B^r. The interpretation of the unary relation symbol R in $\mathbf{B}'$ is the set of r-tuples

$$(b_1^1, \dots, b_{r_1}^1, \dots, b_1^s, \dots, b_{r_s}^s)$$

such that $(b_1^i, \dots, b_{r_i}^i) \in R_i^{\mathbf{B}}$ for every $i \in \{1, \dots, s\}$. The interpretation of the binary relation symbol S in $\mathbf{B}'$ is the arc relation of the De Bruin graph on the universe of B^r. In other words, the pair of r-tuples

$$((b_1, \dots, b_r), (b_1', \dots, b_r'))$$

belongs to $S^{\mathbf{B}'}$ if and only if $b_i = b_{i+1}'$ for every $i \in \{1, \dots, r-1\}$. This defines $\mathbf{B}'$.

Fact 1 *There exists an r-ary quantifier-free reduction from* CSP$(\mathbf{B})$ *to* CSP$(\mathbf{B}')$.

Proof: An arbitrary σ-structure $\mathbf{A}$ is encoded into a τ-structure $\mathbf{I}(\mathbf{A})$ as follows: the universe of $\mathbf{I}(\mathbf{A})$ is defined by the formula $\varphi_U(\mathbf{x}) = \mathrm{true}$, where $\mathbf{x} = (x^1, \dots, x^r)$. The relation R is defined by the formula $\varphi_R(\mathbf{x}_1, \dots, \mathbf{x}_s) = R_1(\mathbf{x}_1) \wedge \cdots \wedge R_s(\mathbf{x}_s)$, where $\mathbf{x}_i = (x_i^1, \dots, x_i^{r_i})$. And the relation S is defined by the formula $\varphi_S(x_1, \dots, x_r, x_1', \dots, x_r')$ below:

$$x_1 = x_2' \wedge x_2 = x_3' \wedge \cdots \wedge x_{r-1} = x_r'.$$

We need to check that $\mathbf{A} \to \mathbf{B}$ if and only of $\mathbf{I}(\mathbf{A}) \to \mathbf{B}'$. Given a homomorphism $h : \mathbf{A} \to \mathbf{B}$, define $g : \mathbf{I}(\mathbf{A}) \to \mathbf{B}'$ by $g((a_1, \dots, a_r)) = (h(a_1), \dots, h(a_r))$, which gives a homomorphism. Given a homomorphism $h : \mathbf{I}(\mathbf{A}) \to \mathbf{B}'$, first use the shift relation S to show that for every $a \in A$ there exists $b(a) \in B$ such that $h((a, \dots, a)) = (b(a), \dots, b(a))$ and $h((a_1, \dots, a_r)) = (b(a_1), \dots, b(a_r))$. Now note that this $b : \mathbf{A} \to \mathbf{B}$ is a homomorphism. □

Now we transform $\mathbf{B}'$ into a digraph $\mathbf{H}$. Before we proceed, we need some terminology about paths in directed graphs. A 1 zig-zag path starting at u and ending at v is an arc from u to v. A 0 zig-zag path starting at u and ending at v is an arc from v to u. If x is a string of zeros and ones, then an $x1$ zig-zag path starting at u and ending at v is an x zig-zag path starting at u and ending at some w from which there is an arc to v. A $x0$ zig-zag path starting at u and ending at v is

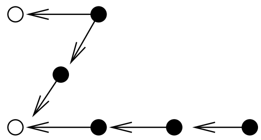

Figure 1: A fragment of $\mathbf{H}$ corresponding to two points from the original $\mathbf{B}'$. These appear in $\mathbf{H}$ too, here as void circles that we call b_1 (top) and b_2 (bottom). There is a 011 zig-zag starting at b_1 and ending at b_2 and a 111 zig-zag ending at b_2. This means that $(b_1, b_2) \in S^{\mathbf{B}'}$ and $b_2 \in R^{\mathbf{B}'}$. $\diamond$

an x zig-zag path starting at u and ending at some w to which there is an arc from v.

We are back to the construction of $\mathbf{H}$. The set of vertices of $\mathbf{H}$ is B' together with three new vertices for every $b \in R^{\mathbf{B}}$ and two new vertices for every pair $(b, b') \in S^{\mathbf{B}}$. The three vertices for every $b \in R^{\mathbf{B}}$ are used to form a 111 zig-zag path starting at any of these vertices and ending at b. The two vertices for every pair $(b, b') \in S^{\mathbf{B}}$ are used to form a 011 zig-zag path starting at b and ending at b'. Figure 1 illustrates this construction.

Fact 2 *There exists a ternary quantifier-free reduction with* 11 *parameters from* CSP($\mathbf{B}'$) *to* CSP($\mathbf{H}$).

Proof: An arbitrary (R, S)-structure $\mathbf{A}'$ is encoded into a digraph $\mathbf{I}(\mathbf{A}')$ as follows. For every point a in $\mathbf{A}'$ we add a new 11011 zig-zag path starting at a new vertex, formed by new vertices, and ending at a. For every point a in $R^{\mathbf{A}'}$ we add a new 111 path starting at a new vertex and ending at a. And for every pair (a, a') in $S^{\mathbf{B}'}$ we add a new 011 zig-zag path starting at a and ending at a'. Figure 2 illustrates this construction.

Let $\mathbf{y} = (a_0, b_1, \ldots, b_5, c_1, \ldots, c_3, d_1, d_2)$ be a collection of eleven parameters (variables). In the quantifier-free reduction we use (a, a, a_0) to encode the copy of a, and $(a, a, b_1), \ldots, (a, a, b_5)$ to endode the 11011 zig-zag path that ends at (a, a, a_0). We also use $(a, a, c_1), \ldots, (a, a, c_3)$ to encode the 111 zig-zag path that ends at (a, a, a_0) for $a \in R^{\mathbf{A}'}$, and $(a, a', d_1), (a, a', d_2)$ to encode the 011 zig-zag path that starts at (a, a, a_0) and ends at (a, a', a_0) for $(a, a') \in S^{\mathbf{A}'}$. Thus, the universe of $\mathbf{I}(\mathbf{A}', \mathbf{y})$ is given by the formula $\varphi_U(x, x', z, \mathbf{y})$ below:

$$\begin{aligned} & (x = x' \wedge (z = a_0 \vee z = b_1 \vee z = b_2 \vee z = b_3 \vee z = b_4 \vee z = b_5)) \\ \vee \quad & (x = x' \wedge R(x) \wedge (z = c_1 \vee z = c_2 \vee z = c_3)) \\ \vee \quad & (S(x, x') \wedge (z = d_1 \vee z = d_2)). \end{aligned}$$

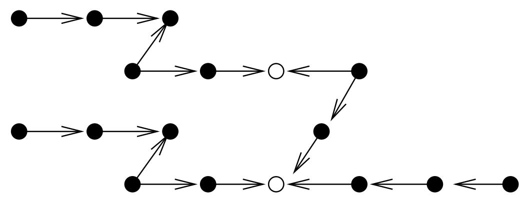

Figure 2: A fragment of $\mathbf{I}(\mathbf{A}')$ corresponding to two points from the original $\mathbf{A}'$. These appear in $\mathbf{I}(\mathbf{A}')$ too, here as void circles that we call a_1 (top) and a_2 (bottom). There are two 11011 zig-zags ending at a_1 and a_2, respectively. There is also a 011 zig-zag starting at a_1 and ending at a_2, and a 111 zig-zag ending at a_2. This means that $(a_1, a_2) \in S^{\mathbf{A}'}$ and $a_2 \in R^{\mathbf{A}'}$. $\diamond$

The arc relation of $\mathbf{I}(\mathbf{A}', \mathbf{y})$ is given by the formula $\varphi_E(x_1, x_1', z_1, x_2, x_2', z_2, \mathbf{y})$ below:

$$\begin{aligned}
& (S(x_1, x_1') \wedge x_2 = x_2' = x_1 \wedge z_1 z_2 = d_1 a_0) \\
\vee \quad & (S(x_1, x_1') \wedge x_2 = x_1 \wedge x_2' = x_1' \wedge z_1 z_2 = d_1 d_2) \\
\vee \quad & (S(x_1, x_1') \wedge x_2 = x_2' = x_1' \wedge z_1 z_2 = d_2 a_0) \\
\vee \quad & (x_1 = x_1' = x_2 = x_2' \wedge (z_1 z_2 = b_1 b_2 \vee z_1 z_2 = b_2 b_3 \vee \\
& \vee z_1 z_2 = b_4 b_3 \vee z_1 z_2 = b_4 b_5 \vee z_1 z_2 = b_5 a_0)) \\
\vee \quad & (x_1 = x_1' = x_2 = x_2' \wedge R(x_1) \wedge (z_1 z_2 = c_1 c_2 \vee z_1 z_2 = c_2 c_3 \vee \\
& \vee z_1 z_2 = c_3 a_0)).
\end{aligned}$$

We are left to prove that $\mathbf{A}' \to \mathbf{B}'$ if and only if $\mathbf{I}(\mathbf{A}', \mathbf{y}) \to \mathbf{H}$. Given a homomorphism $h : \mathbf{A}' \to \mathbf{B}'$, define $g((a, a, a_0)) = h(a)$. Then extend g to map the 11011 zig-zag path that ends in (a, a, a_0) to two consecutive 011 zig-zag paths in $\mathbf{H}$, the second ending in $h(a)$. Note that such paths exist in $\mathbf{H}$ by the definition of the shift relation in $\mathbf{B}'$. For every $a \in R^{\mathbf{A}'}$ we have $h(a) \in R^{\mathbf{B}'}$, so we map the 111 zig-zag path that ends at (a, a, a_0) to the 111 zig-zag path that ends at $h(a)$ in $\mathbf{H}$. For every $(a, a') \in S^{\mathbf{A}'}$ we have $(h(a), h(a')) \in S^{\mathbf{B}'}$, so we map the 011 zig-zag path that starts at (a, a, a_0) and ends at (a', a', a_0) to the 011 zig-zag path that starts at $h(a)$ and ends at $h(a')$ in $\mathbf{H}$. This gives a homomorphism $g : \mathbf{I}(\mathbf{A}') \to \mathbf{H}$.

Conversely, given a homomorphism $h : \mathbf{I}(\mathbf{A}') \to \mathbf{H}$, use the 11011 zig-zag paths to argue that for every $a \in A'$ there exists a $b(a) \in B'$ such that $h((a, a, a_0)) = b(a)$. Then use the 111 zig-zag paths to argue that the R-relation is preserved, and use the 011 zig-zag paths to argue that the S-relation is preserved. This gives a homomorphism from $\mathbf{A}'$ to $\mathbf{B}'$. $\square$

Composing the two facts we get part (1) in Theorem 7 which is what we were after. For part (2), the idea is to first reduce CSP($\mathbf{H}$) to CSP($\mathbf{B}'$), and then CSP($\mathbf{B}'$) to CSP($\mathbf{B}$). The first reduction, as described in [9], is this.

Start with an arbitrary digraph $\mathbf{G}$ and first check that $\mathbf{G}$ is acyclic and does not contain paths of length 4 or more. Otherwise we know $\mathbf{G} \not\to \mathbf{H}$. After that, iteratively collapse in-neighbors of vertices having at least one in-neighbor and at least one out-neighbor until any such vertex has in-degree exactly one. Repeat with out-neighbors of vertices having at least one in-neighbor and at least one out-neighbor until any such vertex also has out-degree exactly one. Then, collapse neighbors of a point starting a path of length 3 so that every such vertex has out-degree exactly one. For points starting at least two paths of length 2, collapse the out-going paths of length 2 into a single path of length 2. For vertices of out-degree two or more, collapse out-neighbors that do not belong to a path of length 2 until every such vertex has at most one out-neighbor that does not belong to a path of length 2. Now we view vertices of out-degree 0 as points a of an (R, S)-structure. Then we impose an R-constraint on every vertex a on which a 111 zig-zag path ends, and an S-constraint on every pair (a, a') with a 011 zig-zag path starting at a and ending at a'. To argue the correctness of this reduction it is important to remember that every vertex in $\mathbf{B}'$ ends two consecutive 011 zig-zag paths by the definition of the shift relation.

It remains to reduce CSP($\mathbf{B}'$) to CSP($\mathbf{B}$). But this is in fact easier. Given an arbitrary instance $\mathbf{A}'$ of CSP($\mathbf{B}'$), replace every point of $\mathbf{A}'$ by r copies, then collapse those that should be equal as dictated by $S^{\mathbf{A}'}$ as if it were a partial shift relation, and finally add the tuples to the relations as dictated by $R^{\mathbf{A}'}$.

Remark Let us point out that the existence of quantifier-free reductions between CSPs as in Theorem 7 cannot be taken for granted. For example, Dalhaus [5] noted that CNF-SAT does not reduce to 3-SAT under quantifier-free reductions. However, CNF-SAT is not quite a CSP in the form we consider in this paper as it has unbounded arity. Another more appropriate example is the following. Consider the two problems HORN-3-SAT and AFFINE-3-SAT. More concretely, HORN-3-SAT is the CSP specified by the template on $\{0, 1\}$ for 3-clauses of the two possible Horn types $(x \vee \overline{y} \vee \overline{z})$ or $(\overline{x} \vee \overline{y} \vee \overline{z})$. Similarly, AFFINE-3-SAT is the CSP specified by the template on $\mathrm{GF}(2)$ for the equations on three variables of the two possible types: $x + y + z = 0$ and $x + y + z = 1$. These two problems are solvable in polynomial-time, and in fact, the first is $\mathbf{P}$-complete under logspace reductions. Therefore, AFFINE-3-SAT reduces to HORN-3-SAT under logspace reductions, yet the reduction cannot be made quantifier-free. Indeed, it is known that HORN-3-SAT has bounded width while AFFINE-3-SAT does not (see [9] and Section 6). As a result, there cannot be a quantifier-free reduction from AFFINE-3-SAT to HORN-3-SAT by Lemma 1. This simple example shows how the strong claim of quantifier-freeness in Theorem 7 cannot be taken for granted without check.

4.2 Consequences of the quantifier-free reduction

Let us now spell out the consequences of Theorem 7. We mentioned already that AFFINE-3-SAT is a CSP that is known not to have bounded width, but is tractable since we can run Gaussian elimination in polynomial time. Let $\mathbf{H}$ be the digraph given by Theorem 7 for the template of AFFINE-3-SAT. Then CSP$(\mathbf{H})$ polynomially reduces to AFFINE-3-SAT, so it is tractable, but AFFINE-3-SAT reduces to CSP$(\mathbf{H})$ by quantifier-free reductions, so CSP$(\mathbf{H})$ does not have bounded width.

Corollary 1 *There exists an acyclic digraph* $\mathbf{H}$ *with* 368 *vertices and* 432 *edges, maximum degree* 2*, and height* 3 *such that* CSP$(\mathbf{H})$ *is tractable but does not have bounded treewidth duality.*

Let us verify that the number of vertices and edges of $\mathbf{H}$ is correct. The template of AFFINE-3-SAT has two points in its universe, and two ternary relations with four tuples each. This gives a $\mathbf{B}'$ with $2^6 = 64$ elements and 16 tuples in R. In $\mathbf{H}$, we have one copy of each of the 64 points. We also have 3 points and 3 edges for every tuple, which makes 48 additional points and edges. And finally 2 points and 3 edges for every pair of 6-tuples in $\{0,1\}^6$ that are related by the shift relation. There are $2^7 = 128$ such pairs, so we get 256 new points and 384 new edges. In total we have 368 points and 432 edges.

The second consequence is about triads. A triad is an oriented tree with a single node of degree 3. It was shown by Hell, Nesetril and Zhu [13] that there exists a triad $\mathbf{T}$ such that CSP$(\mathbf{T})$ is NP-complete. Consequently, CSP$(\mathbf{T})$ does not have bounded width unless P = NP. A close look at [13] reveals that the NP-hardness of CSP$(\mathbf{T})$ is proved by exhibiting a quantifier-free reduction from NOT-ALL-EQUAL-3-SAT, which is the CSP with the template below:

$$(\{0,1\}, \{(0,0,1),(0,1,0),(1,0,0),(0,1,1),(1,0,1),(1,1,0)\}).$$

Thus, it suffices to show that NOT-ALL-EQUAL-3-SAT does not have bounded width. This follows from the following more or less folklore result (see [9] for the definition of the class SNP called Strict NP and see [5] for related results).

Theorem 8 *The problems* 3-SAT *and* NOT-ALL-EQUAL-3-SAT *are SNP-complete under quantifier-free reductions.*

Proof sketch: Every SNP-sentence can be written in the form

$$(\exists X_1)\cdots(\exists X_k)(\forall x_1)\cdots(\forall x_t)\phi,$$

with ϕ a conjunction of formulas of the form

$$\psi(\mathbf{x}) \vee \bigvee_{j\in I} X_j(\mathbf{x}_{\pi_j}) \vee \bigvee_{j\in J} \neg X_j(\mathbf{x}_{\pi_j}), \tag{6}$$

where ψ is a quantifier-free formula without X_j's. By factoring out, we may assume that no pair of such formulas have the same X_j's part. By introducing existentially quantified auxiliary second-order variables if necessary, we may assume as well that each such formula has exactly three occurrences of X_j's. The quantifier-free reduction to 3-SAT is now clear: We define a new propositional variable $Y_{j,\mathbf{a}_\pi}$ for each atomic formula of the form $X_j(\mathbf{x}_\pi)$ and every t-tuple $\mathbf{a}$. Then, for every conjunct in ϕ as in (6) and every t-tuple $\mathbf{a}$, we put a clause

$$\bigvee_{j\in I} Y_{j,\mathbf{a}_{\pi_j}} \vee \bigvee_{j\in J} \neg Y_{j,\mathbf{a}_{\pi_j}}. \tag{7}$$

in the 3-CNF formula if and only if $\psi(\mathbf{a})$ holds.

For NOT-ALL-EQUAL-3-SAT it suffices to show that there is a quantifier-free reduction from 3-SAT. This follows from the seminal work of Schaeffer [26]. Indeed, Schaeffer proved that every relation R_i in the template of 3-SAT is definable in the template of NOT-ALL-EQUAL-3-SAT by means of an existential-positive first-order formula of the form

$$\varphi_i(x,y,z) = (\exists x_1)\cdots(\exists x_k)\phi_i(x,y,z,x_1,\ldots,x_k),$$

where ϕ_i is a conjunction of atomic formulas. The reduction is now clear: besides the original variables of the 3-SAT instance, we create k variables $x_1,\ldots,x_k$ for every tuple (x,y,z) from the 3-SAT instance. Then we place the tuples in the atomic formulas in $\phi_i(x,y,z,x_1,\ldots,x_k)$ in the NOT-ALL-EQUAL-3-SAT instance, for every $(x,y,z)\in R_i$ from the 3-SAT instance. □

Since AFFINE-3-SAT is a CSP, it belongs to SNP [9], so there is a quantifier-free reduction from AFFINE-3-SAT to NOT-ALL-EQUAL-3-SAT, so the latter does not have bounded width.

With all the above:

Corollary 2 *There is a triad* $\mathbf{T}$ *such that* CSP($\mathbf{T}$) *does not have bounded treewidth duality.*

The power of the logical approach stems from the fact that we do not need to painfully analyze homomorphisms from arbitrary structures of treewidth k to complicated transformed instances and templates of the problems. Definability issues alone already tell us that bounded width is ruled out. Let us insist that these results would not have been possible without the recent result in [8] (Theorem 6 above) about the disposal of negations. This is because the quantifier-free reductions require negations. But strangely enough, only inequalities are needed in the quantifier-free reductions and only to state that the parameters are pairwise distinct. In any case, we do not see a way around this, and we think it would be painful, if possible at all, to prove these results without Theorem 6.

5 Application: Finite dualities and first-order logic

Recall from Section 3 the definition of finitary duality. It is easy to see that CSP$(\mathbf{B})$ has finitary duality if and only if CSP$(\mathbf{B})$ is definable in $\neg\exists$FO. But when is CSP$(\mathbf{B})$ definable in full first-order logic FO? This is the question addressed in this section.

A priori FO is much more expressive than $\exists$FO since, in particular, the properties expressible in the latter are always closed under homomorphisms. But what if we restrict ourselves to properties that are closed under homomorphisms? Is every property that is closed under homomorphisms and definable in FO also definable in $\exists$FO? This question is known as the "Homomorphism Preservation Conjecture" which was a central open problem in finite model theory for years and is now a Theorem thanks to Rossman [25]. Here we are able to verify the conjecture for the particular case of CSP$(\mathbf{B})$ problems, whose complements are always closed under homomorphisms, as it is easy to see. Our proof, which appeared in the conference version of this paper in preliminary form, is based entirely on locality arguments for first-order logic and is thus independent of [25].

Theorem 9 *Let σ be a relational vocabulary and let $\mathbf{B}$ be a σ-structure. Then, the following are equivalent:*

1. CSP$(\mathbf{B})$ *is definable in* FO*,*
2. CSP$(\mathbf{B})$ *is definable in* $\neg\exists$FO*.*

The proof of this result relies on a density result for minimal models of first-order sentences that are closed under homomorphisms. This was first established by Ajtai and Gurevich [1] as an interesting application of Gaifman's Local Lemma. Before we state the density result we need to recall some definitions. Recall that $\mathbf{A}$ is a *minimal model of* φ if $\mathbf{A}$ is a model of φ and every proper substructure of $\mathbf{A}$ is not a model of φ. Similarly, recall that the *Gaifman graph of* $\mathbf{A}$, denoted by $\mathcal{G}(\mathbf{A})$, is the graph whose set of vertices is A and whose edges relate every pair of distinct elements that appear together in some tuple of $\mathbf{A}$. If $\mathbf{G} = (V, E)$ is a graph and $S \subseteq V$ is a subset of its vertices, we say that S is *d-scattered in* $\mathbf{G}$ if $d_{\mathbf{G}}(u, v) \geq d$ for every pair of distinct elements u, v in S, where $d_{\mathbf{G}}(u, v)$ denotes the length of the shortest path between u and v in $\mathbf{G}$.

It will also be useful to introduce the concept of the Gaifman hypergraph of a σ-structure, which we denote by $\mathcal{H}(\mathbf{A})$. This is the hypergraph whose vertices are the points in A, and whose hyperedges are the sets of points that form tuples in $\mathbf{A}$. Since we can define distances in the hypergraph in a natural way, namely, as lengths of shortest paths, a d-scattered set in a hypergraph is defined then analogously to the graph case. Note that distances in the Gaifman graph $\mathcal{G}(\mathbf{A})$ and in the Gaifman hypergraph $\mathcal{H}(\mathbf{A})$ of a σ-structure $\mathbf{A}$ coincide, so a d-scattered set in $\mathcal{G}(\mathbf{A})$ is also a d-scattered set in $\mathcal{H}(\mathbf{A})$ and vice-versa. This allows us to state the result by Ajtai and Gurevich in terms of the Gaifman hypergraph, which is the form we need later on.

Theorem 10 ([1]) *Let φ be a first-order sentence such that the class of its finite models is closed under homomorphisms. Then, for every $s \geq 0$, there exist integers $d \geq 0$ and $m \geq 0$ such that if $\mathbf{A}$ is a finite minimal model of φ, then there is no $B \subseteq A$ of size at most s such that $\mathcal{H}(\mathbf{A} - B)$ has a d-scattered set of size m.*

Here we use the notation $\mathbf{A} - B$ to denote the substructure of $\mathbf{A}$ induced by the set $A - B$, where A is the universe of $\mathbf{A}$. For a hypergraph $\mathbf{H} = (V, E)$ and a set of vertices B we use a similar notation $\mathbf{H} - B$ to denote the hypergraph with vertices $V - B$ and edges $\{e : e \in E, e \subseteq V - B\}$. Note that in some texts, $\mathbf{H} - B$ is used to denote the hypergraph with vertices $V - B$ and edges $\{e - B : e \in E\}$. We do not mean that. Note that with our notation we have $\mathcal{H}(\mathbf{A} - B) = \mathcal{H}(\mathbf{A}) - B$.

We will use Theorem 10 in connection with the easy fact proved below. In the conference version of this paper, we proved it only for graphs even though we needed it for hypergraphs. Here we extend the proof to hypergraphs, which is only a bit more delicate. Recall the definition of girth of a hypergraph from Section 2.

Lemma 2 *For every d, m and r, there exists a g and an n such that if $\mathbf{H} = (V, E)$ is a hypergraph with at least n points, edges of size at most r, and girth at least g, then there exists a vertex $u \in V$ such that $\mathbf{H} - \{u\}$ contains a d-scattered set of size m.*

Proof: Fix d, m and r. Let $g = (m + 2)(d + 1) + 1$ and $n = (mr)^{(m+2)(d+1)}$. Suppose that $\mathbf{H} = (V, E)$ is a hypergraph with at least n points, edges of size at most r, and girth at least g. We may assume that $\mathbf{H}$ is connected; otherwise add one edge between every pair of components until the hypergraph is connected, and work with the new hypergraph. Clearly, this does not create a scattered set where none existed before. Since the girth of $\mathbf{H}$ is at least three, no pair of edges share more than one point. This will be useful in the following. Now we consider two cases: either $\mathbf{H}$ contains a vertex u with at least m non-loop edges touching it, or not. In the first case, let S be a set of m different neighbors of u, each from a different non-loop edge touching u. Note that such neighbors exist because no pair of edges share more than one point. Since the girth of $\mathbf{H}$ is at least $g > d + 2$, the set S is d-scattered in $\mathbf{H} - \{u\}$. Its size is at least m. Consider now the second case; namely, that every vertex is touched by at most $m - 1$ non-loop edges, of course each with at most r elements. Then, since $\mathbf{H}$ is connected, its girth is at least $g > (m + 2)(d + 1)$, and its size is at least $n = (mr)^{(m+2)(d+1)}$, there exists a simple path of length $(m + 2)(d + 1)$. For every $i \in \{1, \dots, m\}$, let v_i be the $i(d + 1)$-th vertex in the path. Since the girth of $\mathbf{H}$ is at least $g > 2d$, the set $S = \{v_i : 1 \leq i \leq m\}$ is d-scattered in $\mathbf{H}$. Its size is m, so the lemma is proved. $\square$

With Theorem 10 and Lemma 2 in hand, it will suffice to show that every structure $\mathbf{A} \not\to \mathbf{B}$ can be replaced by another structure $\mathbf{A}' \to \mathbf{A}$ whose Gaifman hypergraph has large girth and still $\mathbf{A}' \not\to \mathbf{B}$. The existence of such an $\mathbf{A}'$ can be derived from a probabilistic construction due to Feder and Vardi [9, Theorem 5] [1]:

Theorem 11 (Theorem 5 in [9]) *Let $\mathbf{A}$ and $\mathbf{B}$ be finite σ-structures and let g be an integer. Then, there exists a structure $\mathbf{A}'$ satisfying the following properties:*

1. $\mathbf{A}' \to \mathbf{A}$,
2. $\mathbf{A} \to \mathbf{B}$ *if and only if* $\mathbf{A}' \to \mathbf{B}$,
3. *the girth of $\mathcal{H}(\mathbf{A}')$ is at least g.*

Here we provide an easier, less ad-hoc proof of a weaker result. Our proof is weaker in two respects: the first is that we only work it out for digraphs, and the second is that instead of guaranteeing girth at least g, we can only guarantee odd-girth at least g (the odd-girth is the length of the shortest odd-cycle). Nonetheless, we believe the proof we provide can be useful to facilitate a better understanding of

[1] The proof in [9, Theorem 5] provides an $\mathbf{A}'$ that is polynomially bounded in $\mathbf{A}$. However, for our purposes, the polynomial bound is irrelevant. It is also worth pointing out that, in [9, Theorem 5], clause (3) is stated as "the girth of $\mathbf{A}'$ is at least g". In their terminology, this means that any $g - 1$ different tuples $\mathbf{a}_1, \dots, \mathbf{a}_{g-1}$ in $\mathbf{A}'$ of arities $r_1, \dots, r_{g-1}$ involve at least $1 + \sum_i (r_i - 1)$ points. But this is equivalent to saying that the length of the shortest cycle (if any) is at least g.

some of the key results in [9]. We learned recently that exactly this proof is already in [12] and goes back to [21], and the statement goes under the name of "sparse incomparability lemma". In the course of the proof we will need the fact that there exist graphs of arbitrarily large odd girth and arbitrarily large chromatic number (see Section 2.5 in [12]). Let us state the weak form of Theorem 11 as a corollary and prove it.

Corollary 3 *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be digraphs and let* g *be an integer. Then, there exists a digraph* $\mathbf{A}'$ *satisfying the following properties:*

1. $\mathbf{A}' \to \mathbf{A}$,
2. $\mathbf{A} \to \mathbf{B}$ *if and only if* $\mathbf{A}' \to \mathbf{B}'$,
3. *the odd-girth of* $\mathcal{H}(\mathbf{A}')$ *is at least* g.

Proof: Assume σ contains a single binary relation symbol R. Let m be the cardinality of B, and let n be the cardinality of A. Let $\mathbf{G} = (V, E)$ be a graph whose girth is at least g and whose chromatic number is at least $m^n + 1$. Let $\mathbf{C}$ be the σ-structure defined as follows. The universe C of $\mathbf{C}$ is V. The interpretation of R in $\mathbf{C}$ is:

$$R^{\mathbf{C}} = \{(u, v) \in V^r : \{u, v\} \in E\}.$$

Note that the Gaifman hypergraph and the Gaifman graph of $\mathbf{C}$ coincide, and $\mathcal{G}(\mathbf{C})$ is precisely $\mathbf{G}$.

Now we define $\mathbf{A}' = \mathbf{C} \times \mathbf{A}$. In other words, the universe of $\mathbf{A}'$ is $C \times A$, and the tuple $((c_1, a_1), (c_2, a_2))$ belongs to $R^{\mathbf{A}'}$ if and only if $(c_1, c_2) \in R^{\mathbf{C}}$ and $(a_1, a_2) \in R^{\mathbf{A}}$. We claim that $\mathbf{A}'$ satisfies the three properties of the Theorem when girth is replaced by odd-girth.

Property (1) is clear since the projection mapping shows that $\mathbf{C} \times \mathbf{A} \to \mathbf{A}$. Property (3) is also clear since any odd-cycle in the Gaifman graph of $\mathbf{C} \times \mathbf{A}$ projects into an odd-cycle[2] in the Gaifman graph of $\mathbf{C}$, and the odd-girth of the latter is at least g. Let us now prove property (2). Obviously, if $\mathbf{A} \to \mathbf{B}$, then also $\mathbf{C} \times \mathbf{A} \to \mathbf{B}$ by (1). Conversely, suppose that $h : \mathbf{C} \times \mathbf{A} \to \mathbf{B}$ is a homomorphism. For every $c \in C$ and $a \in A$, let $h_c(a) = h(c, a)$. Observe that each h_c is a mapping from A to B, but not necessarily a homomorphism. However, there are at most m^n mappings from A to B, and since the chromatic number of $\mathbf{G}$ is at least $m^n + 1$, there must exist an edge $\{c, c'\} \in E$ such that $h_c = h_{c'}$; otherwise $\mathbf{G}$ would be m^n-colorable. We claim that $h' = h_c = h_{c'}$ is a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. To see this, note that if $(a_1, a_2) \in R^{\mathbf{A}}$, then $((c, a_1), (c', a_2)) \in R^{\mathbf{A}'}$, so

$$(h((c, a_1)), h((c', a_2))) \in R^{\mathbf{B}}.$$

[2]Critically, only odd-cycles are guaranteed to map to cycles since an even cycle could alternate between the two end-points of an edge of $\mathbf{G}$.

But since $h_c = h_{c'}$, this is precisely $(h'(a_1), h'(a_2))$. This completes the proof of the weak form of Theorem 11. $\square$

Finally, we are ready for the proof of Theorem 9.

Proof of Theorem 9: Suppose that CSP($\mathbf{B}$) is definable by a first-order sentence φ. It suffices to show that $\neg\varphi$ has finitely many minimal models. Suppose the opposite and we will contradict Theorem 10. Fix integers d and m, and let r be the maximum arity of σ. Let g and n be as in Lemma 2. Let $\mathbf{A}$ be a minimal model of $\neg\varphi$ of size at least n. Let $\mathbf{A}'$ be as in Theorem 11. Since $\mathbf{A}$ is a model of $\neg\varphi$, we have $\mathbf{A} \not\rightarrow \mathbf{B}$, so $\mathbf{A}' \not\rightarrow \mathbf{B}$, so $\mathbf{A}'$ is a model of $\neg\varphi$ as well. Let $\mathbf{A}''$ be a substructure of $\mathbf{A}'$ that is a minimal model of $\neg\varphi$. We have $\mathbf{A}'' \rightarrow \mathbf{A}' \rightarrow \mathbf{A}$. The homomorphic image of $\mathbf{A}''$ into $\mathbf{A}$ is a substructure of $\mathbf{A}$. But since $\neg\varphi$ is preserved under homomorphisms and $\mathbf{A}$ is a minimal model of $\neg\varphi$, the homomorphism must be surjective. This shows that $|A''| \geq n$. Note that the girth of $\mathcal{H}(\mathbf{A}'')$ is still at least g since $\mathcal{H}(\mathbf{A}'')$ is a subhypergraph of $\mathcal{H}(\mathbf{A}')$. Now, by Lemma 2, there exists a vertex $u \in A''$ such that $\mathcal{H}(\mathbf{A}'') - \{u\}$ contains a d-scattered set S of size m. Note that $\mathcal{H}(\mathbf{A}'') - \{u\} = \mathcal{H}(\mathbf{A}'' - \{u\})$. This shows that for every d and m there exist some minimal model of $\neg\varphi$ for which removing a single element produces a d-scattered set of size m. This contradicts Theorem 10. $\square$

Theorem 9, together with the obvious fact that finitary duality and definability in $\neg\exists$FO are the same, shows that those CSP($\mathbf{B}$) that are first-order definable are precisely those having finitary duality. On the other hand, Nešetřil and Tardif [22] (see also [20, Theorem 3.13]) characterized exactly those CSP($\mathbf{B}$) problems having finitary duality through an explicit identification of the singleton dualities. In the following, a *σ-strictree* is a σ-structure whose *shadow* is a tree, where the *shadow* of a σ-structure $\mathbf{A}$ is the multigraph with set of vertices A and having an edge between a and b if there exists a tuple $\mathbf{a} = (a_1, \ldots, a_r)$ in $\mathbf{A}$ such that $a = a_i$ and $b = a_{i+1}$ for some $i \in \{1, \ldots, r-1\}$. Note that if the shadow is a tree then it does not have loops or parallel edges.

It was shown in [22] that for every σ-strictree $\mathbf{T}$, there exists a σ-structure $\mathbf{D}(\mathbf{T})$ that is the dual of $\mathbf{T}$; in other words, $\mathbf{A} \rightarrow \mathbf{D}(\mathbf{T})$ if and only if $\mathbf{T} \not\rightarrow \mathbf{A}$. Moreover, $\mathbf{D}(\mathbf{T})$ is explicitly defined from $\mathbf{T}$. Now we get our corollary:

Corollary 4 (to Theorem 9 and to [22]) *Let $\mathbf{B}$ be a finite σ-structure. Then, the following are equivalent:*

1. CSP($\mathbf{B}$) *is definable in* FO*,*
2. CSP($\mathbf{B}$) *is definable in* $\neg\exists$FO*,*

3. CSP$(\mathbf{B})$ *has finitary duality,*
4. *there is a finite collection of σ-strictrees $\{\mathbf{T}_1, \ldots, \mathbf{T}_t\}$ that is a duality for* CSP$(\mathbf{B})$ *and such that $\mathbf{B}$ is homomorphically equivalent to $\prod_{i=1}^{t} \mathbf{D}(\mathbf{T}_i)$.*

Let us point out that the strong clause (4) in Corollary 4 implies that there is a semi-decision procedure to tell, given $\mathbf{B}$, whether CSP$(\mathbf{B})$ is first-order definable. Indeed, the construction of $\mathbf{D}(\mathbf{T})$ is computable given $\mathbf{T}$, so it suffices to find the finite set of σ-strictrees $\{\mathbf{T}_1, \ldots, \mathbf{T}_t\}$ and check that $\mathbf{B}$ is homomorphically equivalent to $\prod_{i=1}^{t} \mathbf{D}(\mathbf{T}_i)$. Note, however, that there is no a priori bound on this set, so this is only a semi-decision procedure.

The point is that this is already a non-trivial result since the obvious phrasings of the statement "CSP$(\mathbf{B})$ is first-order definable" are in Σ_2, the second level of the arithmetic hierarchy. Indeed, the most obvious phrasing that CSP$(\mathbf{B})$ is first-order definable is that there exists a first-order formula φ such that for every instance $\mathbf{A}$ we have that $\mathbf{A} \models \varphi$ iff $\mathbf{A} \to \mathbf{B}$. A second less obvious phrasing can be stated in terms of quantifier ranks and Ehrenfeucht-Fraïssé games, but the complexity does not change and remains Σ_2. Corollary 4, instead, gives a Σ_1 statement which, to our knowledge, was not known before the conference version of this paper appeared. All this notwithstanding, it is interesting to comment here that the recent work by Larose, Loten and Tardif [18] has finally shown that it is decidable, given $\mathbf{B}$, whether CSP$(\mathbf{B})$ has finitary duality. Hence, it is also decidable whether CSP$(\mathbf{B})$ is first-order definable by Corollary 4. Indeed, the problem is NP-complete, which completely settles the issue.

Let us also point out that for vocabularies of maximum arity two, as in digraphs, σ-strictrees have treewidth one. Therefore, by Theorem 4, a finite obstruction set for CSP$(\mathbf{B})$ that consists of σ-strictrees implies definability in the two-variable fragment $\neg\exists\mathrm{FO}^2$ of $\neg\exists\mathrm{FO}$. Thus, first-order logic collapses to $\neg\exists\mathrm{FO}^2$ for digraph CSPs.

6 Application: Width Bounds

How does one prove that a certain CSP$(\mathbf{B})$ does not have bounded width? This question was addressed already in the original paper by Feder and Vardi [9] where two different approaches were suggested. The first approach consists in reducing the question of proving width lower bounds to that of proving lower bounds for the size of monotone Boolean circuits. Indeed, if CSP$(\mathbf{B})$ has bounded width, then its complement is definable in Datalog, which easily implies that its complement can be decided by polynomial-size monotone Boolean circuits. Feder and Vardi made use of this observation to show that CSP$(\mathbf{B})$'s with the so-called *ability to count* do not have bounded width. Razborov's celebrated result [24] proving monotone circuit lower bounds for deciding if a bipartite graph has a perfect matching came in handy. The second suggested approach consists in reducing the question of proving width lower bounds to that of designing winning strategies in certain two-player pebble game. As we understand it today, this game is the existential pebble game of Section 2 as shown in Theorem 5. This is the approach we take here.

We will show in this section that AFFINE-3-SAT does not have bounded width and that, in fact, the required width is a tight function of the treewidth of the instance. The lower bound is, to our knowledge, new.

6.1 Discussion

A first idea to get width lower bounds via pebbles games would be to use the methods developed for studying the complexity of propositional resolution and the so-called Tseitin formulas. The techniques employed in that area would show that if a graph $\mathbf{G}$ is a sufficiently good *expander*, then the system of equations corresponding to the Tseitin formula of $\mathbf{G}$ requires large width to refute as satisfiable. The results that follow will provide a tighter version of this since we show that large treewidth, as opposed to large expansion, is already enough to guarantee large refutation width of the Tseitin systems.

For completeness, let us remind the reader that the concept of expansion of a graph $\mathbf{G} = (V, E)$ that is used in the context of resolution is measured by the so-called *edge-expansion* of $\mathbf{G}$: as

$$e(\mathbf{G}) = \min\{E(A, V - A) : A \subseteq V, 1/3|V| \leq |A| \leq 2/3|V|\},$$

where $E(A, B)$ is the number of edges with an endpoint in A and the other in B. Using the fact that graphs of small treewidth have small balanced separators, it can be shown that if a sufficiently large graph has treewidth at most k, then its edge-expansion is at most k^2. Thus, bounded treewidth implies bounded edge-expansion, but the converse need not be true. Take for example the disjoint union of two $n \times n$-grids; the edge-expansion is 0 but the treewidth is n. If we do not like disconnected graphs, we can also take two disjoint $n \times n$-grids joined by an edge; in that case the edge-expansion is 1 and the treewidth is again unbounded.

6.2 Tseitin systems

Let us define now the Tseitin system of a given graph $\mathbf{G}$. Fix $k > 0$. Let $\mathbf{G} = (V, E)$ be a 3-regular connected graph with n vertices. Let $u_0 \in V$ be a distinguished vertex. The Tseitin system of $(\mathbf{G}, u_0)$, denoted by $\mathbf{A} = \mathbf{A}(\mathbf{G}, u_0)$, contains one variable for every edge $e \in E$, so $A = \{x_e : e \in E\}$, and one equation

$$\sum_{u \in e} x_e = m(u)$$

for every vertex $u \in V$, where $m(u) = 1$ if $u = u_0$ and $m(u) = 0$ otherwise.

From now on, let $\mathbf{B}$ be the template structure of AFFINE-3-SAT, that is, $\mathbf{B}$ has $\{0, 1\}$ as universe and two ternary relations R_0 and R_1, where R_d is interpreted as $\{(a, b, c) : a + b + c = d\}$ where addition is modulo two. The first claim is straightforward: Tseitin systems are always unsatisfiable.

Claim 1 $\mathbf{A} \not\to \mathbf{B}$.

Proof: If we add all equations together we obtain

$$2 \sum_{e \in E} x_e = \sum_{u \in V} m(u) = m(u_0) = 1.$$

Note that the left-hand side is even but the right-hand side is odd. This shows that $\mathbf{A}$ is unsatisfiable. □

Next we claim that if the treewidth of $\mathbf{G}$ is k, then the Duplicator wins the existential $(\lfloor k/2 \rfloor - 1)$-pebble game on $\mathbf{A}$ and $\mathbf{B}$. In fact, we will also prove that this is tight up to constant factors. Let us start with the upper bound and leave the main result of this section as *finale*.

6.3 The upper bound

Again, $\mathbf{A} = \mathbf{A}(\mathbf{G}, u_0)$ is the Tseitin system of a 3-regular connected graph $\mathbf{G}$. What is the Gaifman graph $\mathcal{G}(\mathbf{A})$ of $\mathbf{A}$ and how does it relate to $\mathbf{G}$? As it happens, $\mathcal{G}(\mathbf{A})$ is exactly the line graph of $\mathbf{G}$. Recall that the line graph of $\mathbf{G}$ is the graph whose vertex set is E, and whose edges are the pairs of edges of $\mathbf{G}$ that share exactly one vertex. Moreover, the treewidth of the line graph can be bounded by the treewidth of $\mathbf{G}$ times the maximum degree:

Lemma 3 *Let $\mathbf{G}$ be a graph of maximum degree d. If the treewidth of $\mathbf{G}$ is k, then the treewidth of the line graph of $\mathbf{G}$ is at most $d(k + 1) - 1$.*

Proof: Let (T, L) be a tree-decomposition of $\mathbf{G}$ of width k. Let L' be the following alternative labelling of T: for every $t \in T$, let $L'(t) = \{e : L(t) \cap e \neq \emptyset\}$. In other words, $L'(t)$ is the set of edges that are incident to some vertex in $L(t)$. Note that the size of each $L'(t)$ is bounded by $d(k + 1)$. We claim that (T, L') is a tree-decomposition of the line graph $\mathbf{L}$ of $\mathbf{G}$. Since the two endpoints of any edge of $\mathbf{G}$

appear in some bag of (T, L), it is clear that every edge of $\mathbf{L}$ also appears in some bag of (T, L'). For the connectivity condition, suppose that the edge $e = \{u, v\}$ of $\mathbf{G}$, which is a vertex of $\mathbf{L}$, appears in both $L'(t)$ and $L'(t')$. If one of the endpoints of e appears also in both $L(t)$ and $L(t')$, then this same endpoint appears in every bag of (T, L) in the path from t to t'. In that case, e appears also in every bag of (T, L') in this path. Suppose now that u appears only in $L(t)$ and v only in $L(t')$. Since e is an edge of $\mathbf{G}$, there must exist at least one bag $L(t'')$ containing both endpoints of e. But then, by the connectivity condition of (T, L), one of these t'' must appear in the path between t and t'. It follows that every bag of (T, L) in this path contains either u or v, so every bag of (T, L') in this path contains e. This concludes the proof. □

With this lemma in hand, it suffices to use the fact first noted in [17] that if an instance of a CSP has treewidth k, then the existential $(k+1)$-pebble game decides homomorphism. We have all necessary material to give a short proof. Recall that in this section $\mathbf{A} = \mathbf{A}(\mathbf{G}, u_0)$ is the Tseitin system of a 3-regular connected graph.

Lemma 4 *If k is the treewidth of $\mathbf{G}$, then the Spoiler wins the existential $3(k+1)$-pebble game on $\mathbf{A}$ and $\mathbf{B}$.*

Proof: If the treewidth of $\mathbf{G}$ is k, then the treewidth of the line graph of $\mathbf{G}$ is at most $3(k+1) - 1$ by Lemma 3. But the line graph of $\mathbf{G}$ is precisely the Gaifman graph of $\mathbf{A}$, so the treewidth of $\mathbf{A}$ is at most $3(k+1) - 1$. Of course we have $\mathbf{A} \rightarrow \mathbf{A}$ and $\mathbf{A} \not\rightarrow \mathbf{B}$. So apply Theorem 4 and get a $\wedge\exists\mathrm{FO}^{3(k+1)}$-sentence ψ that holds in $\mathbf{A}$ and not in $\mathbf{B}$. It follows from Theorem 2 that $\mathbf{A} \not\leq^{k+1}_{\infty\omega} \mathbf{B}$. □

6.4 The lower bound

We will need the fact that the treewidth of a graph is characterized by the k-Cops-and-Robber game [27]. In the k-Cops-and-Robber game there are k Cops and one Robber. The Robber stands at a vertex of the graph and can run at great speed to any other vertex along a path of the graph, but may not run through or to a vertex containing a Cop. Each of the k Cops is either on a vertex or in a helicopter. The objective of the Robber is to escape from the Cops by moving to a different vertex each time he sees that a Cop is approaching and wants to land in his vertex. If he has a strategy to escape forever, we say that the Robber wins. A formal definition of the game can be found in [27].

Theorem 12 ([27]) *Let $\mathbf{G}$ be a graph and let k be an integer. Then, the following are equivalent:*

1. *the treewidth of* $\mathbf{G}$ *is at least* k*,*
2. *the Robber wins the* k*-Cops-and-Robber game on* $\mathbf{G}$.

Now we are ready to state and prove the main claim of this section. Compare with Lemma 4.

Lemma 5 *If* k *is the treewidth of* $\mathbf{G}$*, then the Duplicator wins the existential* $(\lfloor k/2 \rfloor - 1)$*-pebble game on* $\mathbf{A}$ *and* $\mathbf{B}$.

Proof: In a nutshell, the strategy of the Duplicator is to maintain a walk in $\mathbf{G}$ starting at u_0 and ending at some v_0 for which none of the three edges that are incident to v_0 are pebbled. With such a path in hand, the Duplicator will be safe if she sets $e \mapsto 1$ for edges that appear an odd number of times in the walk, and $e \mapsto 0$ for edges that appear an even number of times in the walk. Naturally, the last vertex of the walk and the walk itself will change dinamically during the play of the game using a strategy in the Cops-and-Robber game on $\mathbf{G}$ that is played on the side. Details follow.

Let P be a walk in $\mathbf{G}$ starting at u_0 and ending at $v_0 \neq u_0$. We define a map h_P from the edges of $\mathbf{G}$ to $\{0, 1\}$ by setting $h_P(e) = 1$ for every e that appears an odd number of times in P, and $h_P(e) = 0$ for every e that appears an even number of times in P. We note that, for every set of edges C that are not incident to v_0, the map h_P restricted to C is a partial homomorphism from $\mathbf{A}$ to $\mathbf{B}$. Indeed, in a walk, every internal vertex belongs to an even number of edges of the walk, while the extreme vertices u_0 and v_0, by the fact they are different, belong to an odd number of edges of the walk.

We are ready to define the strategy for the Duplicator. Let P be the walk starting at u_0 currently held by the Duplicator. Initially, we let P be simply any edge incident to u_0. Let v_0 be the last vertex of the walk P. Let C be the set of pebbled edges at the begining of the i-th round of the game and let D be the set of vertices that are incident to these edges. The Duplicator will maintain the invariant that the set of vertices having a Cop in the side game on $\mathbf{G}$ is precisely $D \cup \{u_0\}$. Note that $|D| \leq 2|C| + 1$.

Suppose now that $|C| \leq k/2 - 2$ and that the Spoiler places an unused pebble on edge e. If $e \in C$, let the Duplicator repeat its previous move on e and proceed to the next round of the game with the same P. Let us assume now that $e \notin C$. If e is not incident to v_0, the last vertex of P, we let the Duplicator reply according to h_P. Since the restriction of h_P to $C \cup \{e\}$ is still a partial homomorphism, the game can proceed to the next round. Suppose now that e is incident to v_0. Now, in the Robber-Cops game played on the side, place two Cops over the vertices u and v forming the edge e. Recall that $|D| \leq 2|C| + 1$ and $|C| \leq k/2 - 2$, so there are enough Cops to proceed. Before the Cop lands in v_0, the Robber can escape through a path Q avoiding $D \cup \{u_0\} \cup \{u, v\}$, to a new vertex v_0'. Let P' be the walk that goes from u_0 to v_0 as in P and then from v_0 to v_0' as in Q. Finally, let the Duplicator reply according to $h_{P'}$. Notice that h_P and $h_{P'}$ agree on C since Q avoids all pebbled edges. The new walk kept by the Duplicator for the next round is P'.

It remains to see how to proceed when the Spoiler removes a pebble from some $e \in C$. In this case the Duplicator removes the corresponding pebble from $\mathbf{T}$ and updates the configuration of the Robber-Cops game by removing the Cops from the vertices that form the edge e, except for u_0 which always keeps a guarding Cop. The invariant is maintained, and this completes the proof. □

Take now a collection of 3-regular graphs of unbounded treewidth. An interesting example is the class of (toroidal) brick graphs. This shows that AFFINE-3-SAT requires unbounded width even on the Tseitin systems of this very particular class of graphs.

Corollary 5 ([9]) AFFINE-3-SAT *does not have bounded width.*

This follows from Lemma 5 and clause (6) in Theorem 5. The same argument would work to show that AFFINE-4-SAT requires large width on the Tseitin systems corresponding to the class of (toroidal) grid graphs. Note that such grids are 4-regular, and of course, have unbounded treewidth.

Acknowledgement We thank Phokion Kolaitis, Jarik Nešetřil, and the referees of LICS 2005 for useful comments.

References

[1] M. Ajtai and Y. Gurevich. Datalog vs. first-order logic. *Journal of Computer and System Sciences*, 49(3):562–588, 1994.

[2] A. Atserias. On sufficient conditions for unsatisfiability of random formulas. *Journal of the ACM*, 51(2):281–311, 2004. Preliminary version in LICS 2002.

[3] A. Atserias, A. Dawar, and Ph. G. Kolaitis. On preservation under homomorphisms and unions of conjunctive queries. In *31st Annual ACM-SIGACT-SIGACT Symposium on Principles of Database Systems*, 2004.

[4] A. Bulatov. Malt'sev constraints are tractable. Technical report PRG-RR-02-05, Oxford University, 2002.

[5] E. Dalhaus. Reduction to NP-complete problems by interpretations. In E. Börger, Rödding, and Hasenjaeger, editors, *Logic and Machines: Decision Problems and Complexity*, volume 171 of *Lecture Notes in Computer Science*, pages 357–365. Springer-Verlag, 1984.

[6] V. Dalmau, Ph. G. Kolaitis, and M. Y. Vardi. Constraint satisfaction, bounded treewidth, and finite variable logics. In *8th International Conference on Principles and Practice of Constraint Programming (CP)*, volume 2470 of *Lecture Notes in Computer Science*, pages 310–326. Springer, 2002.

[7] V. Dalmau, A. Krokhin, and B. Larose. First-order retraction problems for posets and reflexive graphs. In *19th IEEE Symposium on Logic in Computer Science*, 2004.

[8] T. Feder and M. Vardi. Homomorphism closed vs. existential positive. In *18th IEEE Symposium on Logic in Computer Science*, pages 311–320, 2003.

[9] T. Feder and M. Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory. *SIAM Journal of Computing*, 28(1):57–104, 1998.

[10] W. Gutjahr, W. Welzl, and G. Woeginger. Polynomial graph-colorings. *Discrete Applied Mathematics*, 35:29–45, 1992.

[11] P. Hell and J. Nešetřil. On the complexity of H-coloring. *Journal of Combinatorial Theory*, 48:92–119, 1990.

[12] P. Hell and J. Nešetřil. *Graph Homomorphisms*. Oxford University Press, 2003.

[13] P. Hell, J. Nešetřil, and X. Zhu. Complexity of tree homomorphisms. *Discrete Applied Mathematics*, 70:23–36, 1996.

[14] P. Hell, J. Nešetřil, and X. Zhu. Duality and polynomial testing of tree homomorphisms. *Transactions of the American Mathematical Society*, 348(4):1281–1297, 1996.

[15] P. Jeavons, D. Cohen, and M. Gyssens. Closure properties of constraints. *Journal of the ACM*, 44(4):527–548, 1997.

[16] Ph. G. Kolaitis and M. Y. Vardi. On the expressive power of Datalog: tools and a case study. *Journal of Computer and System Sciences*, 51:110–134, 1995.

[17] Ph. G. Kolaitis and M. Y. Vardi. Conjunctive-query containment and constraint satisfaction. *Journal of Computer and System Sciences*, 61(2):302–332, 2000.

[18] B. Larose, C. Loten, and C. Tardif. A characterisation of first-order constraint satisfaction problems. In *21st IEEE Symposium on Logic in Computer Science*, 2006.

[19] B. Larose and L. Zádori. Bounded width problems and algebras. *Algebra Universalis*, 2006. to appear.

[20] J. Nešetřil. Aspects of structural combinatorics (graph homomorphisms and their use). *Taiwanese Journal of Mathematics*, 3:381–423, 1999.

[21] J. Nešetřil and V. Rödl. Chromatically optimal rigid graphs. *Journal of Combinatorial Theory Series B*, 46:133–141, 1989.

[22] J. Nešetřil and C. Tardif. Duality theorems for finite structures (characterising gaps and good characterisations). *Journal of Combininatorial Theory Series B*, 80:80–97, 2000.

[23] J. Nešetřil and Z. Zhu. On bounded treewidth duality of graphs. *Journal of Graph Theory*, 23:151–162, 1996.

[24] A. A. Razborov. Lower bounds on monotone complexity of the logical permanent. *Matematicheskie Zametki*, 37(6):887–900, 1985.

[25] B. Rossman. Existential positive types and preservation under homomorphisms. In *20th IEEE Symposium on Logic in Computer Science*, 2005.

[26] T. J. Schaefer. The complexity of satisfability problems. In *10th Annual ACM Symposium on the Theory of Computing*, pages 216–226, 1978.

[27] P. D. Seymour and R. Thomas. Graph searching and a min-max theorem for tree-width. *Journal of Combinatorial Theory, Series B*, 58:22–33, 1993.