Semantic Width and the Fixed-Parameter Tractability of Constraint Satisfaction Problems

Hubie Chen[1] , **Georg Gottlob**[2,3] , **Matthias Lanzinger**[3] and **Reinhard Pichler**[3]

[1]Birkbeck, University of London
[2]Oxford University
[3]TU Wien

h.chen@dcs.bbk.ac.uk, georg.gottlob@cs.ox.ac.uk, {mlanzing, pichler}@dbai.tuwien.ac.at

Abstract

Constraint satisfaction problems (CSPs) are an important formal framework for the uniform treatment of various prominent AI tasks, e.g., coloring or scheduling problems. Solving CSPs is, in general, known to be NP-complete and fixed-parameter intractable when parameterized by their constraint scopes. We give a characterization of those classes of CSPs for which the problem becomes fixed-parameter tractable. Our characterization significantly increases the utility of the CSP framework by making it possible to decide the fixed-parameter tractability of problems via their CSP formulations. We further extend our characterization to the evaluation of unions of conjunctive queries, a fundamental problem in databases. Furthermore, we provide some new insight on the frontier of PTIME solvability of CSPs. In particular, we observe that bounded fractional hypertree width is more general than bounded hypertree width only for classes that exhibit a certain type of exponential growth. The presented work resolves a long-standing open problem and yields powerful new tools for complexity research in AI and database theory.

1 Introduction

CSPs are a fundamental problem of artificial intelligence. As a unifying formal framework, they play a foundational role in many areas of AI research, see e.g., [Kumar, 1992; Narváez, 2018]. However, the unifying aspect of CSPs has not yet reached its full potential. While a CSP formulation of a problem allows for reuse of common algorithmic strategies and implementations [Gottlob *et al.*, 2000; Do and Kambhampati, 2001], results in computational complexity still often require individual investigation, with little help from the framework. A complexity characterization for CSP would allow researchers to finally leverage the CSP framework also for strong computational complexity results, hence greatly simplifying the study of all problems that can be formulated as CSPs. The consequences and wide-reaching applications of such a characterization motivate our central research question. Note that throughout this paper, the pa-

rameterized complexity of CSPs always refers to the problem parameterized by the size of its constraint scopes.

Research Challenge: Is there a natural characterization of the fixed-parameter tractable classes of CSPs?

To be precise, we study what is referred to as the uniform CSP problem in the literature. In the uniform problem, we are interested in how the structure of constraint scopes affects the complexity of the problem, i.e., we characterize restrictions to the structure of constraint scopes. In the nonuniform problem, one considers restrictions to the constraint relations. Here, in a classic result, [Hell and Nesetril, 1990] gave an elegant characterization of PTIME solvability. More recently, [Bulatov, 2017] and [Zhuk, 2017] were able to independently establish a powerful dichotomy theorem. However, results for the nonuniform case do not translate to the uniform problem.

There is a long line of research devoted to the (parameterized) computational complexity of solving CSPs based on structural parameters of their associated hypergraphs. In a landmark result, [Grohe, 2007] resolved the question for a restricted class of CSPs; namely those with bounded arity. There, PTIME decidability is fully characterized by bounded treewidth modulo homomorphic equivalence. Moreover, for bounded arity, we have fixed-parameter tractability if and only if the problem is solvable in PTIME.

To tackle the problem beyond bounded arity, a number of generalizations of treewidth have been developed that provide sufficient conditions for tractably solving CSPs, the most important of which are hypertree width [Gottlob *et al.*, 2002] and fractional hypertree width [Grohe and Marx, 2014]. Yet, bounding these parameters yields only sufficient conditions for tractability. A necessary condition for unbounded arity remains elusive. In the parameterized space, a highly impressive result by [Marx, 2013] was able to characterize those hypergraphs, i.e., problem structures, that always allow for fixed-parameter tractable evaluation by the *submodular width* of the hypergraphs. However, while this result is closely related to our goal, the fact that the characterization is on the hypergraph level significantly limits its applicability in our setting: When we consider the CSP formulation of a problem, then the complexity of our problem does not depend on the complexity of other, unrelated, CSPs that happen to have the same underlying hypergraphs. Hence, characterizing on

the hypergraph level restricts us to a worst-case that may not be connected to the problem we want to study (this point is discussed in detail in Section 3.1).

Despite their unquestionable importance, Grohe's and Marx's characterizations do not answer our research question. Instead, we introduce a new parameter – *semantic submodular width* (sem-$subw$) – to capture the minimal submodular width over the (infinite) equivalence class of semantically equivalent CSPs. We show that it is still possible to decide sem-$subw$ and find the minimal semantically equivalent CSP in time that depends only on the size of the parameter. Following that, we give a reduction from Marx's setting to ours, which allows us to prove the necessary lower bound. Akin to Marx's characterization, our result assumes the Exponential Time Hypothesis [Impagliazzo *et al.*, 2001]; a standard assumption of parameterized complexity.

Main Result 1: Assuming the Exponential Time Hypothesis, a class of CSPs is fixed-parameter tractable if and only if it has bounded semantic submodular width.

Through the well-known equivalence of CSP to the homomorphism problem as well as conjunctive query containment [Kolaitis and Vardi, 2000] and evaluation [Maier, 1983], our main result also applies to those important problem families. By adapting our notion of sem-$subw$ from CSPs to the more general notion of *unions of conjunctive queries* (UCQs) accordingly, we can also extend our characterization result to UCQs, an important and widely studied class of query languages in database theory [Sagiv and Yannakakis, 1980; Atserias *et al.*, 2006].

Main Result 2: Assuming the Exponential Time Hypothesis, a class of UCQs is fixed-parameter tractable if and only if it has bounded semantic submodular width.

With the question of fixed-parameter tractability resolved, we shift our attention to PTIME solvable classes of CSPs. Here, a characterization of tractable restrictions for the uniform CSP problem remains an open question. We briefly discuss how our parameterized results relate to the non-parameterized case. Furthermore, we utilize some recent results on the connection of hypergraph width parameters and Vapnik-Chervonenkis dimension to derive new insight on the frontier of tractability of the uniform CSP problem. In particular, we show that the two most important sufficient conditions for tractable CSP solving – bounded fractional hypertree width and bounded hypertree width – actually collapses for classes of CSPs as long as they do not exhibit a certain kind of, seemingly unnatural, exponential growth.

The rest of the paper is structured as follows. Section 2 recalls necessary definitions for constraint satisfaction problems, unions of conjunctive queries, and relevant hypergraph width parameters. We expand on the differences to Marx's characterization in Section 3.1 before we present our two main results in Sections 3.2 and 3.3. Section 4 presents some new insights regarding the PTIME solvability of CSPs. We end with concluding remarks in Section 5.

Some preliminary results related to this paper were previously published in [Gottlob *et al.*, 2019].

2 Preliminaries

2.1 Parameterized Complexity

Parameterized complexity enables a more fine-grained study of computational complexity. Here, we give an abridged definition of the notions necessary for this paper. For full definitions and details we refer to [Flum and Grohe, 2006].

For an alphabet of symbols Σ, a *parameterized problem* is given as a pair (P, κ) of a problem $P \subseteq \Sigma^*$ and its parameterization κ that maps each string in Σ^* to a parameter.

We say that a parameterized problem (P, κ) is *fixed-parameter tractable* if there exists an algorithm that decides whether a given string $x \in \Sigma^*$ is in P in time $f(\kappa(x)) poly(|x|)$, where f is a computable function and $poly$ is a polynomial.

Let (P, κ) and (P', κ') be two parameterized problems. A *fpt-reduction* from (P, κ) to (P', κ') is a mapping $R : \Sigma^* \to \Sigma^*$ with the following properties:
(1) $x \in P \iff R(x) \in P'$ for every $x \in \Sigma^*$,
(2) R is computable in time $f(\kappa(x)) poly(|x|)$ (f is computable), and
(3) there is a computable function g such that $\kappa'(x) \leq g(\kappa(x))$ for all $x \in \Sigma^*$.
We say (P, κ) is fpt-reducible to (P', κ'), denoted $(P, \kappa) \leq (P', \kappa')$. The class of fixed-parameter tractable problems is closed under fpt-reductions.

Our main results assume the Exponential Time Hypothesis, which states that 3-SAT with n variables can not be decided in $2^{o(n)}$ time [Impagliazzo *et al.*, 2001]. This is a standard assumption of parameterized complexity theory.

2.2 Constraint Satisfaction Problems

We formalize CSPs as a relational homomorphism problem. A *signature* is a finite set of relation symbols with associated arities. A *(relational) structure* $\mathbf{A}$ (over signature σ) consists of a domain A and an interpretation $R^{\mathbf{A}}$ for each relation symbol R in the signature. Let $\mathbf{A}, \mathbf{B}$ be relational structures, we call a function $h : A \to B$ a *homomorphism* from $\mathbf{A}$ into $\mathbf{B}$, if for every relation symbol R and all $(x_1, \ldots, x_{ar(R)}) \in R^{\mathbf{A}}$ also $(h(x_1), \ldots, h(x_{ar(R)})) \in R^{\mathbf{B}}$, where $ar(R)$ is the arity of R. We write $|\mathbf{A}| = |\sigma| + |A| + \sum_{R \in \sigma} |R^{\mathbf{A}}| ar(R)$ for the size of structure $\mathbf{A}$.

We call an ordered pair $(\mathbf{A}, \mathbf{B})$ of structures a *constraint satisfaction problem instance*. Intuitively, $\mathbf{A}$ expresses the constraint scopes and $\mathbf{B}$ the permitted assignments for each constraint scope. For a class $\mathcal{A}$ of structures, the corresponding *constraint satisfaction problem*, denoted $\text{CSP}(\mathcal{A})$, is the following decision problem.

$\text{CSP}(\mathcal{A})$	
Instance:	A CSP instance $(\mathbf{A}, \mathbf{B})$ where $\mathbf{A} \in \mathcal{A}$.
Question:	Is there a homomorphism from $\mathbf{A}$ into $\mathbf{B}$?

By slight abuse of notation, we also call $\mathcal{A}$ a class of *constraint satisfaction problems*. Note that what we call $\text{CSP}(\mathcal{A})$ is sometimes denoted as $\text{CSP}(\mathcal{A}, -)$ to emphasize that we are dealing with the uniform CSP problem (cf., [Grohe, 2007]). Furthermore, constants play no role in our considerations since they can be eliminated by straightforward preprocessing.

A *hypergraph* H is a tuple $(V(H), E(H))$, where $V(H)$ is the set of vertices and $E(H) \subseteq 2^{V(H)}$ the set of *hyper-edges*. For a set $U \subseteq V(H)$, we define the subhypergraph induced by U as $H[U] = (U, E')$ where $E' = \{e \cap U \mid e \in E(H)\} \setminus \{\emptyset\}$. The hypergraph $H(\mathbf{A})$ of a structure $\mathbf{A}$ is the hypergraph where the vertices equal A and $e \in E(H(\mathbf{A}))$ if and only if there exists some relation symbol R such that some permutation of e is contained in $R^{\mathbf{A}}$. The hypergraph of a CSP instance $(\mathbf{A}, \mathbf{B})$ is the hypergraph of $\mathbf{A}$, i.e., the hypergraph of a CSP instance represents only the structure of its constraint scopes. We are interested in how this structure affects the complexity of the decision problem. We thus consider the CSP decision problem parameterized by its constraint scope structure:

p-CSP$(\mathcal{A})$	
Instance:	A CSP instance $(\mathbf{A}, \mathbf{B})$ where $\mathbf{A} \in \mathcal{A}$.
Parameter:	$\|\mathbf{A}\|$
Question:	Is there a homomorphism from $\mathbf{A}$ into $\mathbf{B}$?

For a class $\mathcal{H}$ of hypergraphs, let $Struct[\mathcal{H}]$ denote all structures whose hypergraphs are in $\mathcal{H}$. We will abbreviate the problem CSP$(Struct[\mathcal{H}])$ to CSP$(\mathcal{H})$, i.e., CSP restricted to those instances whose hypergraphs are in $\mathcal{H}$. The analogue applies to p-CSP.

For two structures $\mathbf{A}$ and $\mathbf{A}'$, we say $\mathbf{A}$ is *homomorphically equivalent* to $\mathbf{A}'$, or $\mathbf{A} \simeq \mathbf{A}'$, if there exists a homomorphism from $\mathbf{A}$ into $\mathbf{A}'$ and vice versa. The *core* of a structure $\mathbf{A}$, denoted $core(\mathbf{A})$, is the minimal structure (with regards to the number of tuples) that is homomorphically equivalent to $\mathbf{A}$. It is not hard to verify that every structure has a unique (up to isomorphism) core. For a class of structures $\mathcal{A}$, we write $core(\mathcal{A})$ for the class of cores of structures in $\mathcal{A}$.

In the context of CSPs, we say a structure $\mathbf{A}$ is *contained* in structure $\mathbf{A}'$ if for every $\mathbf{B}$ we have that if $(\mathbf{A}, \mathbf{B})$ has a solution, then $(\mathbf{A}', \mathbf{B})$ also has a solution. It is easy to see that $\mathbf{A}$ is contained in $\mathbf{A}'$ if and only if there exists a homomorphism from $\mathbf{A}'$ to $\mathbf{A}$. If two structures $\mathbf{A}$ and $\mathbf{A}'$ are contained within each other, we say that they are *semantically equivalent* (we write $\mathbf{A} \equiv \mathbf{A}'$). Hence, if $\mathbf{A} \equiv \mathbf{A}'$ then for every $\mathbf{B}$ we have that $(\mathbf{A}, \mathbf{B})$ has a solution if and only if $(\mathbf{A}', \mathbf{B})$ has a solution. Furthermore, note that $\mathbf{A} \equiv \mathbf{A}'$ if and only if $\mathbf{A} \simeq \mathbf{A}'$, i.e., homomorphic equivalence equals semantic equivalence. In particular, $(\mathbf{A}, \mathbf{B})$ is always equivalent to $(core(\mathbf{A}), \mathbf{B})$.

2.3 Unions of Conjunctive Queries

Please note that, for consistency and brevity, we will define unions of conjunctive queries via CSPs. This does not match the standard presentations of the problem but is equivalent to them.

An instance of the *(boolean) unions of conjunctive queries* (UCQ) problem is a set of structures $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$, we write $\bigcup_{i=1}^{n} \mathbf{A}_i$, and a structure $\mathbf{B}$ which is usually referred to as the database. We say an instance of the UCQ problem $(\bigcup_{i=1}^{n} \mathbf{A}_i, \mathbf{B})$ has a solution if any of the CSP instances $(\mathbf{A}_i, \mathbf{B})$, for $1 \leq i \leq n$, has a solution. Hence, the accompanying parameterized decision problem for a class of UCQs $\mathcal{U}$ is the following

p-UCQ$(\mathcal{U})$	
Instance:	A UCQ $U = \bigcup_{i=1}^{n} \mathbf{A}_i$ where $U \in \mathcal{U}$ and a database $\mathbf{B}$.
Parameter:	$\sum_{i=1}^{n} \|\mathbf{A}_i\|$
Question:	Does $U, \mathbf{B}$ have a solution?

Analogue to CSPs, the equivalence of UCQs will be important. We say that two UCQs $U = \bigcup_{i=1}^{n} \mathbf{A}_i$ and $U' = \bigcup_{i=1}^{m} \mathbf{A}'_i$ are *semantically equivalent* (we write $U \equiv U'$) if for every structure $\mathbf{B}$, $(U, \mathbf{B})$ has a solution if and only if $(U', \mathbf{B})$ has a solution.

A UCQ $\bigcup_{i=1}^{n} \mathbf{A}_i$ is *non-redundant* if there are no $\mathbf{A}_i$ and $\mathbf{A}_j$ $(i \neq j)$ such that $\mathbf{A}_i$ is contained in $\mathbf{A}_j$. Note that every UCQ can be transformed into an equivalent non-redundant UCQ by repeated deletion of structures that are contained by other structure in the UCQ [Sagiv and Yannakakis, 1980]. We write $nr(U)$ for the UCQ obtained by applying this procedure to make an UCQ U non-redundant. Importantly, as the procedure only deletes structures we have $nr(U) \subseteq U$.

2.4 Decompositions and Their Widths

In this work we will only consider width notions that are based on tree decompositions. A tuple $(T, (B_u)_{u \in T})$ is a *tree decomposition* of a hypergraph H if T is a tree, every B_u is a subset of $V(H)$ and the following two conditions are satisfied:

(1) For every $e \in E(H)$ there is a node $u \in T$ s.t. $e \subseteq B_u$, and

(2) for every vertex $v \in V(H)$, $\{u \in T \mid v \in B_u\}$ is connected in T.

For functions $f : 2^{V(H)} \to \mathbb{R}^+$, the *$f$-width* of a tree decomposition is $\sup\{f(B_u) \mid u \in T\}$ and the f-width of a hypergraph is the minimal f-width over all its tree decompositions. Let $\mathcal{F}$ be a class of functions from subsets of $V(H)$ to the non-negative reals, then the $\mathcal{F}$-width of H is $\sup\{f\text{-width}(H) \mid f \in \mathcal{F}\}$. All such widths are implicitly extended to structures and CSP instances by taking the width of their respective hypergraphs.

The following properties of functions $f : 2^{V(H)} \to \mathbb{R}^+$ are important:

- f is *monotone* if $X \subseteq Y$ implies $f(X) \leq f(Y)$.
- f is called *edge-dominated* if $f(e) \leq 1$ for every $e \in E(H)$.
- f is called *submodular* if $f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$ holds for every $X, Y \subseteq V(H)$.

We say a weight function $\gamma : E(H) \to \mathbb{R}^+$ is a *fractional edge cover* of a set $X \subseteq V(H)$ if for every $v \in X$ we have $\sum_{e \in I_v} \gamma(e) \geq 1$ where I_v is the set of all edges incident to v. If we restrict the co-domain to $\{0, 1\}$, we obtain the definition of an *integral edge cover*. We refer to the total weight $\sum_{e \in E(H)} \gamma(e)$ of an edge cover as the *size* of the edge cover.

For $X \subseteq V(H)$, let $\rho_H(X)$ be the size of the smallest integral edge cover of X by edges in $E(H)$ and $\rho_H^*(X)$ the size of the smallest fractional edge cover of X by edges in $E(H)$. This framework now allows us to define many of the important widths in the current literature.

(Primal) Treewidth of H [Robertson and Seymour, 1986]:
$tw(H) := c\text{-width}$, where $c(X) = \|X\| - 1$.

Generalized hypertree width of H **[Gottlob *et al.*, 2002]:**
$ghw(H) := \rho_H$-width.
Fractional hypertree width of H **[Grohe and Marx, 2014]:**
$fhw(H) := \rho_H^*$-width.
Submodular width of H **[Marx, 2013]:**
$subw(H) := \mathcal{F}$-width(H), where $\mathcal{F}$ is the set of all monotone, edge-dominated, submodular functions b on $2^{V(H)}$ with $b(\emptyset) = 0$.

A notable omission, that is not expressible through this notion of f-width, is *hypertree width (hw)* [Gottlob *et al.*, 2002], which uses the same width function as ghw but imposes an additional restriction on the tree decomposition. Note that these widths spawn a hierarchy in the sense that the following inequality holds for all hypergraphs H:

$$subw(H) \leq fhw(H) \leq ghw(H) \leq hw(H) \leq tw(H) + 1$$

For a class of structures $\mathcal{A}$, we say $\mathcal{A}$ has bounded width if there exists a constant k such that every structure in $\mathcal{A}$ has width $\leq k$. The computational complexity of CSP is tightly linked to this hierarchy of parameters. This connection is summarized by the following two propositions.

Proposition 1 ([Grohe and Marx, 2014]). *Let $\mathcal{C}$ be a class of CSP instances of bounded fhw. Then* CSP$(\mathcal{C})$ *is tractable.*

Proposition 2 ([Marx, 2013]). *Let $\mathcal{H}$ be a recursively enumerable class of hypergraphs. Assuming the Exponential Time Hypothesis, p-CSP$(\mathcal{H})$ is fixed-parameter tractable if and only if $\mathcal{H}$ has bounded submodular width.*

3 Main Results

3.1 Characterization of Hypergraph Classes vs. Classes of CSP Instances

Recall the motivation given in the introduction. Many AI problems have natural CSP formulations and we wish to determine the computational complexity of all such problems through a characterization of the complexity of CSP. In this section we argue why a characterization on the hypergraph level (which ignores relation symbols), as in Proposition 2, is not enough for this goal. The main issue with the hypergraph characterization is that even though a CSP instance may have a highly complex hypergraph structure, it can still be easy to solve. Yet, the complexity of p-CSP$(\mathcal{H})$ expresses only the complexity of the worst-case CSP instances of the given structure. We illustrate this issue in the following example.

Consider the following problem: Given a directed graph G, can we embed (by a homomorphism) a bidirected $n \times n$-grid into G? The corresponding CSP instance $C_n(G) = (\mathbf{A}, \mathbf{B})$ has a single relation symbol E and $\mathbf{B} = G$. As domain of $\mathbf{A}$ we take $\{x_{i,j} \mid i \in [n], j \in [n]\}$ and $E^{\mathbf{A}}$ contains exactly the following tuples specifying the $n \times n$-grid: $(x_{i,j}, x_{i+1,j}), (x_{i+1,j}, x_{i,j})$ for $i \in [n-1], j \in [n]$ and $(x_{i,j}, x_{i,j+1}), (x_{i,j+1}, x_{i,j})$ for $i \in [n], j \in [n-1]$.

We now consider the class $\mathcal{C}$ of all CSP instances $C_n(G)$ for $n \geq 1$ and all graphs G. The hypergraphs of $\mathcal{C}$ are, by definition, exactly the class of $n \times n$-grid graphs $\mathcal{G}_{n \times n}$, which is well-known to have unbounded treewidth [Robertson and Seymour, 1986]. In general, it is difficult to determine the submodular width of graphs since the definition depends on

a supremum over an infinite class of functions. However, Lemma 1 below provides us with a convenient way to recognize that certain classes have unbounded submodular width.

Lemma 1. *Let H be an arbitrary hypergraph and let $rank(H)$ be the maximum edge size in H, then*

$$tw(H) \leq rank(H) \cdot subw(H)$$

Sketch. Let $f : X \mapsto |X|/rank(H)$ be a function on the subsets of $V(H)$. It is easy to verify that f is submodular, edge-dominated and monotone. For any node u of any tree decomposition of H we clearly have $|B_u| = rank(H) \cdot f(B_u)$ and therefore also $tw(H) + 1 = rank(H) \cdot f$-width$(H)$. Since f is submodular, edge-dominated and monotone we also have f-width$(H) \leq subw(H)$ and the statement follows immediately. $\square$

From Lemma 1 we can conclude that $\mathcal{G}_{n \times n}$ also has unbounded submodular width. From Proposition 2 we can thus only deduce that p-CSP$(\mathcal{G}_{n \times n})$ is not fixed-parameter tractable.

However, for every $C_n(G) = (\mathbf{A}, \mathbf{B})$, we have that $core(\mathbf{A})$ is the structure with domain $\{x_1, x_2\}$ and $E^{core(\mathbf{A})} = \{(x_1, x_2), (x_2, x_1)\}$. This is easy to verify, e.g., by observing that an undirected $n \times n$-grid is 2-colorable. Clearly, $(core(\mathbf{A}), \mathbf{B})$ is solvable in polynomial time and it is equivalent to $(\mathbf{A}, \mathbf{B})$. It follows that p-CSP$(\mathcal{C})$ is in fact fixed-parameter tractable (and indeed tractable), despite the complexity of p-CSP$(\mathcal{G}_{n \times n})$. We see that a hypergraph level characterization has inherent shortcomings in establishing lower bounds for specific problem classes.

3.2 Constraint Satisfaction Problems

In this section we prove our characterization theorem for CSPs. The discussion in Section 3.1 shows that unbounded submodular width can still allow for fixed-parameter tractable CSP solving. Hence, we require a new, more general, property to fully capture fixed-parameter tractability. We follow [Barceló *et al.*, 2017] who introduced the notion of *semantic (generalized) hypertree width* and define the following general notion of *semantic widths* of CSPs.

Definition 1. *Let $\mathcal{A}$ be the class of all structures and $w : \mathcal{A} \to \mathbb{R}^+$ be an isomorphism invariant. We define semantic w as $\text{sem-}w(\mathbf{A}) := \inf\{w(\mathbf{A}') \mid \mathbf{A}' \equiv \mathbf{A}\}$.*

Using this definition, we are now ready to state our first main result. We show that the characterization from Proposition 2 can indeed be strengthened to the following characterization of the fixed-parameter tractability of CSP instances.

Theorem 1. *Let $\mathcal{A}$ be a recursively enumerable class of CSPs. Assuming the Exponential Time Hypothesis, p-CSP$(\mathcal{A})$ is fixed-parameter tractable if and only if $\mathcal{A}$ has bounded semantic submodular width.*

Our proof of the theorem relies on two central lemmas. First, we show how bounded semantic submodular width leads to fixed-parameter tractability. The basic idea is simple, instead of solving a CSP instance with possibly arbitrarily high submodular width, we want to solve an equivalent instance with low width. However, it is not clear how to find

such an equivalent instance and whether finding it is decidable. For generalized hypertree width [Barceló *et al.*, 2017] have recently shown, that for any structure $\mathbf{A}$, sem-$ghw(\mathbf{A})$ is precisely $ghw(core(\mathbf{A}))$. Indeed, we show in Lemma 2, that the same connection also holds for the more complex cases of fractional hypertree width and submodular width. Note that for treewidth this property is trivial since treewidth is hereditary, i.e., removing edges from a hypergraph can not increase its treewidth. The width functions considered here are not hereditary and involve additional technical considerations beyond those necessary for the ghw case.

Lemma 2. *For every structure $\mathbf{A}$:*

1. sem-$\rho^*(\mathbf{A}) = \rho^*(core(\mathbf{A}))$
2. sem-$fhw(\mathbf{A}) = fhw(core(\mathbf{A}))$
3. sem-$subw(\mathbf{A}) = subw(core(\mathbf{A}))$

Proof (Sketch). First, since call equivalent structures have isomorphic cores it is enough to show $w(core(\mathbf{A})) \leq w(\mathbf{A})$ to establish that sem-$w(\mathbf{A}) = w(core(\mathbf{A}))$ for any invariant w.

Let H be the hypergraph of $\mathbf{A}$ and H' the hypergraph of $core(\mathbf{A})$. Note that there exists an homomorphism h from $\mathbf{A}$ to $core(\mathbf{A})$ where $h(a) = a$ for all elements in the domain of $core(\mathbf{A})$. From this we can then show that for every tree decomposition $(T, (B_u)_{u \in T})$ of H, there exists a tree decomposition $(T, (B'_u)_{u \in T})$ of H' where $B'_u = B_u \cap V(H')$.

For the fhw case we then use an observation on how fractional edge covers behave under homomorphisms to show that this transformation does not increase the ρ^*-width. Hence, we can transform the tree decomposition for H with minimal ρ^*-width into a new tree decomposition for H' with less or equal ρ^*-width, i.e., $fhw(H') \leq fhw(H)$. The observation for edge covers under homomorphisms also leads to the result for sem-ρ^*.

The $subw$ case requires additional considerations as the width is now defined over a whole class of functions $\mathcal{F}$. We show that for every edge-dominated, submodular function f' over H' there exists an edge-dominated submodular function f over H such that f'-width$(H') \leq f$-width(H). In particular, for every f' this the function $f : X \mapsto f'(X \cap V(H'))$ has the required properties. $\square$

While sem-ρ^* and sem-fhw are less general than sem-$subw$ they will be of further interest in the discussion of PTIME solvability of CSPs in Section 4. In the context of our main result, the most important consequence of Lemma 2 is that we are always able to find the equivalent structure with minimal submodular width by simply computing the core. In principle, finding the core of a structure is intractable (formally, deciding if a structure $\mathbf{A}'$ is the core of a structure $\mathbf{A}$, is DP-complete [Fagin *et al.*, 2005]). However, in our parameterized setting the computation of the core of $\mathbf{A}$ only depends on the parameter.

To establish a lower bound for classes with unbounded semantic submodular width we will make use of previous results from [Chen and Müller, 2015]. A step in our reduction will require an additional definition that helps us fix the domains of individual elements in the reduction. For a structure

$\mathbf{A}$, let $\mathbf{A}^*$ be the expansion of $\mathbf{A}$ by a new fresh unary relation symbol U_a with $U_a^{\mathbf{A}} = \{a\}$ for every element of the domain $a \in A$. For a class of structures $\mathcal{A}$ we write $\mathcal{A}^*$ for $\{\mathbf{A}^* \mid \mathbf{A} \in \mathcal{A}\}$. Our intention is to establish our lower bound by reduction from the hypergraph setting of Proposition 2. We will make use of the following two reductions.

Proposition 3 ([Chen and Müller, 2015]). *Let $\mathcal{A}$ be a recursively enumerable class of structures. Then*

$$p\text{-}\mathrm{CSP}(core(\mathcal{A})^*) \leq p\text{-}\mathrm{CSP}(\mathcal{A})$$

Lemma 3. *Let $\mathcal{A}$ be a recursively enumerable class of structures and let $\mathcal{H}^{\mathcal{A}}$ be the class of hypergraphs of $\mathcal{A}$.*

$$p\text{-}\mathrm{CSP}(\mathcal{H}^{\mathcal{A}}) \leq p\text{-}\mathrm{CSP}(\mathcal{A}^*)$$

Proof. Let $(\mathbf{C}, \mathbf{D})$ be an instance of $\mathrm{CSP}(\mathcal{H}^{\mathcal{A}})$ and let H be the hypergraph of $\mathbf{C}$ and D be the domain of $\mathbf{D}$. Recall that edges can represent multiple constraint scopes, i.e., multiple tuples in $\mathbf{C}$. For each edge $e \in E(H)$, we consider the sets $F_{t_1}, \ldots, F_{t_k}$ of satisfying assignments $e \to D$ for each of the tuples $t_1, \ldots, t_k$ of $\mathbf{C}$ that become edge e in the hypergraph. We then produce the set $F_e = \bigcap_{i=1}^{k} F_{t_i}$ of satisfying assignments over all the tuples for e. Observe that computing F_e for all $e \in E(H)$ is possible in polynomial time.

By definition there exists a structure $\mathbf{A}^*$ in $\mathcal{A}^*$ where $\mathbf{A}$ has hypergraph H. We can compute such a $\mathbf{A}^*$ by enumeration of $\mathcal{A}$ until we find an $\mathbf{A}$ with a matching hypergraph and then computing $\mathbf{A}^*$ from $\mathbf{A}$.

We will reduce $(\mathbf{C}, \mathbf{D})$ to $(\mathbf{A}^*, \mathbf{B})$ where $\mathbf{B}$ is constructed as follows. As the domain of $\mathbf{B}$ we take $A \times D$. For each $a \in A$ we have a $U_a^{\mathbf{A}}$ with $U_a^{\mathbf{A}} = \{a\}$. Let $U_a^{\mathbf{B}} = \{(a, d) \mid d \in D\}$. For each other relation symbol R of $\mathbf{A}$ and each tuple $(a_1, \ldots, a_k) \in R^{\mathbf{A}}$, we add tuples $((a_1, f(a_1)), \ldots, (a_k, f(a_k)))$ to $R^{\mathbf{B}}$ where $f \in F_e$ and e is the hyperedge $\{a_1, \ldots, a_k\}$.

We now show that $(\mathbf{C}, \mathbf{D})$ has a solution iff $(\mathbf{A}^*, \mathbf{B})$ has a solution. First, suppose h is a homomorphism from $\mathbf{C}$ to $\mathbf{D}$ and note that $\mathbf{A}^*$ and $\mathbf{C}$ have the same domain since $\mathbf{A}$ and $\mathbf{C}$ have the same underlying hypergraph. It is then not difficult to see that $g : a \mapsto (a, h(a))$ is a homomorphism from $\mathbf{A}^*$ to $\mathbf{B}$: For the unary relations $U_a^{\mathbf{A}}$, the image trivially exists in $U_a^{\mathbf{B}}$. For the other relations, it is enough to observe hat for every edge e of H, the assignment h restricted to variables in e must be in F_e.

For the other side, observe that a homomorphism g from $\mathbf{A}^*$ to $\mathbf{B}$ must be of the form $a \mapsto (a, h(a))$. We argue that h is a homomorphism from $\mathbf{C}$ to $\mathbf{D}$. As $\mathbf{A}$ and $\mathbf{C}$ have the same domain, h also applies to the domain of $\mathbf{C}$. By definition of F_e we have that for every tuple $\bar{x}$ in $R^{\mathbf{C}}$, h maps to a tuple in $R^{\mathbf{D}}$ as long as $\bar{x}$ is covered by some edge of $\mathbf{A}$. Since the hypergraphs are the same, this holds for all the tuples in $\mathbf{C}$ and therefore h is a homomorphism. $\square$

Proof of Theorem 1. Let $\mathcal{H}^{core(\mathcal{A})}$ be the class of hypergraphs of the structures in $core(\mathcal{A})$. We claim that the two problems $p\text{-}\mathrm{CSP}(\mathcal{A})$ and $p\text{-}\mathrm{CSP}(\mathcal{H}^{core(\mathcal{A})})$ are fpt-reducible to each other. If the claim holds, $p\text{-}\mathrm{CSP}(\mathcal{A})$ is fixed-parameter tractable iff $p\text{-}\mathrm{CSP}(\mathcal{H}^{core(\mathcal{A})})$ is fixed-parameter tractable. By Proposition 2 this is the case iff $\mathcal{H}^{core(\mathcal{A})}$ has

bounded submodular width. By Lemma 2, this is equivalent to $\mathcal{A}$ having bounded semantic submodular width.

What is left, is to show the claim. First, we observe:

$$p\text{-CSP}(\mathcal{A}) \leq p\text{-CSP}(core(\mathcal{A})) \leq p\text{-CSP}(\mathcal{H}^{core(\mathcal{A})})$$

The left reduction holds because $(\mathbf{A}, \mathbf{B})$ is equivalent to $(core(\mathbf{A}), \mathbf{B})$ and computing the core is feasible in $f(|\mathbf{A}|)$ time. The right reduction is trivial since all instances of $p\text{-CSP}(core(\mathcal{A}))$ are also instances of $p\text{-CSP}(\mathcal{H}^{core(\mathcal{A})})$. For the other direction we get the intended reduction by straightforward combination of Lemma 3 and Proposition 3:

$$p\text{-CSP}(\mathcal{H}^{core(\mathcal{A})}) \leq p\text{-CSP}(core(\mathcal{A})^*) \leq p\text{-CSP}(\mathcal{A})$$

$\square$

3.3 Unions of Conjunctive Queries

We now extend the characterization in Theorem 1 from CSPs to UCQs. To do so we first need to introduce a way to extend the relevant definitions to UCQs. For our width notions the natural extension to UCQs is through the maximum of its parts, i.e., for width function w and UCQ $U = \bigcup_{i=1}^{n} \mathbf{A}_i$ let $w(U) := \max\{w(\mathbf{A}_i) \mid i \in [n]\}$. Semantic width functions are defined the same as for CSPs, i.e., $\text{sem-}w := \inf\{w(U') \mid U' \equiv U\}$. However, equivalence of UCQs is more complex than equivalence in CSPs. In particular, the characterization by homomorphic equivalence is no longer applicable. Therefore, some additional effort is required to determine the analogue of Lemma 2. Using the following classic result by Sagiv and Yannakakis we can derive the fitting Lemma 4.

Proposition 4 ([Sagiv and Yannakakis, 1980]). *Let $U = \bigcup_{i=1}^{n} \mathbf{A}_i$ and $U' = \bigcup_{j=1}^{m} \mathbf{A}'_j$ be non-redundant UCQs. Then $U \equiv U'$ if and only if for every $\mathbf{A}_i$ there is a unique $\mathbf{A}'_j$ such that $\mathbf{A}_i \equiv \mathbf{A}'_j$.*

Lemma 4. *Let U be an UCQ, then*

$$\text{sem-}subw(U) = \max\{subw(core(\mathbf{A}_i)) \mid \mathbf{A}_i \in nr(U)\}$$

Proof. It is clear that the right side of the equality is the $subw$ of an UCQ that is equivalent to U. All that is to show is that this is in fact the minimal $subw$ of an equivalent UCQ. For the sake of brevity we will write core-$subw(nr(U))$ for $\max\{subw(core(\mathbf{A}_i)) \mid \mathbf{A}_i \in nr(U)\}$ in the rest of the argument.

Proof is by contradiction. Suppose there exist a UCQ $V \equiv U$ with $subw(V) < \text{core-}subw(nr(U))$. Since $V \equiv U$, clearly also $nr(V) \equiv nr(U)$. Furthermore, since $nr(V) \subseteq V$ (recall the construction of $nr(V)$) we also have $subw(nr(V)) < \text{core-}subw(nr(U))$. Now, from Proposition 4 we have that for every $\mathbf{B}_i \in nr(V)$, there is an equivalent $\mathbf{A}_j \in nr(U)$. By Lemma 2 it follows that $subw(\mathbf{B}_i) \geq \text{sem-}subw(\mathbf{B}_i) = \text{sem-}subw(\mathbf{A}_j) = subw(core(\mathbf{A}_j))$ for all such combinations of $\mathbf{B}_i$ and $\mathbf{A}_j$. From the definition of $subw$ for UCQs this then gives an immediate contradiction of $subw(nr(V)) < \text{core-}subw(nr(U))$. $\square$

From Lemma 4 it is now easy to see, that for a class of UCQs $\mathcal{U}$ with bounded $\text{sem-}subw$, the $p\text{-UCQ}(\mathcal{U})$ problem is fixed-parameter tractable. For every U in $\mathcal{U}$ we can simply compute $nr(U) = \bigcup_{i=1}^{n} \mathbf{A}_i$ and then solve the CSPs

$(core(\mathbf{A}_i), \mathbf{B})$ individually. In combination with Theorem 1 we see that this procedure is fixed-parameter tractable.

To establish the lower bound, we make use of previous work on the complexity of existential positive logic [Chen, 2014]. The result there is stated in a different setting but a translation is not difficult through the well-known equivalence of solving CSPs and model checking of primitive positive first-order formulas.

Proposition 5 (Theorem 3.2 in [Chen, 2014]). *Let $\mathcal{U}$ be recursively enumerable class of non-redundant UCQs and let $\mathcal{A}$ be the class of all individual structures that make up the UCQs in $\mathcal{U}$. Then $p\text{-CSP}(\mathcal{A}) \leq p\text{-UCQ}(\mathcal{U})$.*

Theorem 2. *Let $\mathcal{U}$ be a recursively enumerable class of UCQs. Assuming the Exponential Time Hypothesis, $p\text{-UCQ}(\mathcal{U})$ is fixed-parameter tractable if and only if $\mathcal{U}$ has bounded semantic submodular width.*

Proof. For the case where $\mathcal{U}$ has bounded semantic submodular width we have already given a fixed-parameter tractable procedure for solving $p\text{-UCQ}(\mathcal{U})$ above. We will establish the lower bound by introducing the class $nr(\mathcal{U}) = \{nr(U) \mid U \in \mathcal{U}\}$ as an intermediate.

Suppose $\mathcal{U}$ has unbounded $\text{sem-}subw$ and let $\mathcal{A}$ be the class of all individual structures that make up the UCQs in $nr(\mathcal{U})$. From Lemma 4 it follows that $nr(\mathcal{U})$ and $\mathcal{A}$ both also have unbounded $\text{sem-}subw$. By Proposition 5 we have $p\text{-CSP}(\mathcal{A}) \leq p\text{-UCQ}(nr(\mathcal{U}))$ and therefore, by Theorem 1, $p\text{-UCQ}(nr(\mathcal{U}))$ can not be fixed-parameter tractable.

To finish the proof we show that $p\text{-UCQ}(nr(\mathcal{U})) \leq p\text{-UCQ}(\mathcal{U})$. The reduction is straightforward, an instance $(U, \mathbf{B})$ of $p\text{-UCQ}(nr(\mathcal{U}))$ is reduced to the instance $(U', \mathbf{B})$ of $p\text{-UCQ}(\mathcal{U})$ where $nr(U') \equiv U$. Such an U' can be found by enumeration of $\mathcal{U}$ in time that only depends on the parameter. Since $nr(U') \equiv U$, the reduction is trivially correct. $\square$

4 On the Plain Tractability of CSPs

A characterization for the plain (non-parameterized) tractability of CSPs remains an open question. Here we wish to highlight two consequences of our work and recent developments regarding the connection of fractional hypertree width and the Vapnik-Chervonenkis (VC) dimension of a hypergraph presented in [Gottlob *et al.*, 2020].

Tractability in natural problem classes. Bounded hypertree width (hw), generalized hypertree width (ghw) and fractional hypertree width (fhw) all represent sufficient conditions for tractable CSP solving, with fhw being the most general such property we know of. It is known that hw is bounded if and only if ghw is bounded [Adler *et al.*, 2007]. Furthermore, there exist classes that exhibit bounded fhw but unbounded hw [Grohe and Marx, 2014]. However, all known hypergraph classes with bounded fhw and unbounded hw involve some form of exponential growth that is unlikely to be present in natural problems. It has remained an open question if this exponential growth is essential for the separation of the two width measures.

Below, we give an answer to this question. The technical details of VC dimension are not important here. Rather we

introduce the notion of *exotic* hypergraph classes, a consequence of unbounded VC dimension, to focus on the exponential character of such classes. We are able to state that this property is indeed intrinsic to the separation of bounded fhw and hw. Alternatively, in the contrapositive, we see that for non-exotic classes, a class has bounded fhw if and only if it has bounded hw. In other words, bounded fhw does not allow for additional tractable cases over bounded hw.

Definition 2. *Let $\mathcal{H}$ be a class of hypergraphs. We say that $\mathcal{H}$ is* exotic *if for every integer $n \geq 1$, there exists a $H \in \mathcal{H}$ with a set of n vertices $U \subseteq V(H)$ such that $H[U]$ has at least $2^n - 1$ distinct edges.*

Theorem 3. *For any class $\mathcal{H}$ of hypergraphs, if $\mathcal{H}$ has unbounded hypertree width and bounded fractional hypertree width then $\mathcal{H}$ is exotic.*

Proof (Sketch). As stated above, exoticness is a consequence of unbounded VC dimension, thus we also have that if $\mathcal{H}$ is not exotic, then $\mathcal{H}$ has bounded VC dimension.

The key observation is then that the integrality gap for fractional edge covers can be bounded by a function of the VC dimension. Hence, under bounded VC dimension the integrality gap is constant. By applying this observation to every bag of a tree decomposition with $fhw \leq k$ we can see that the tree decomposition will also have ghw bounded by some function of k and the VC dimension. Due to space restrictions we refer to the proof Theorem 7.8 in [Gottlob *et al.*, 2020] for details.

In summary, if $\mathcal{H}$ has bounded VC dimension, then $\mathcal{H}$ has bounded fhw iff $\mathcal{H}$ has bounded ghw. Recall from above, that also $ghw(\mathcal{H})$ is bounded iff $hw(\mathcal{H})$ is bounded. Hence, the contrapositive of the implication in the theorem holds. $\square$

We can extend the exotic property from hypergraphs to classes of CSPs in the usual way. Recall, that in the context of CSPs, incident edges in the hypergraph correspond to constraints that involve the variable. Hence, if vertices U have $2^{|U|} - 1$ distinct incident edges in the hypergraph, there exists at least one constraint for every possible combination of the corresponding variables in the CSP. We argue that this situation is highly unnatural and believe that this motivates further study of the complexity of non-exotic classes of CSP.

Semantic width and tractability. In the parameterized setting, it is easy to utilize low semantic width to establish upper-bounds as computing the core requires time only in the parameter. For tractability the situation is more problematic. As noted in Section 3.2, finding the core is intractable. Hence, if we have a class with bounded semantic fractional hypertree width, we know that the problem itself is not difficult, but an efficient solution depends on the hard problem of finding the core. We are caught in an unsatisfactory situation where the origin of the hardness is no longer the actual problem but the concrete formulation.

Part of the issue is that utilizing bounded fhw for polynomial evaluation requires a concrete decomposition with low fhw, which then guides the efficient solution of the CSP. Without knowing the core we cannot compute the appropriate decomposition. For bounded generalized hypertree width, Chen and Dalmau were able to show, that for classes of

bounded ghw, there exists an algorithm for solving CSPs in polynomial time without requiring the explicit computation of a decomposition [Chen and Dalmau, 2005]. Their method indeed remains polynomial if only the semantic generalized hypertree width is bounded. Thus, we are able to lift their result to bounded semantic fractional hypertree width for non-exotic classes.

Corollary 1. *Let $\mathcal{C}$ be a non-exotic class of CSPs with bounded semantic fractional hypertree width. Then $\mathrm{CSP}(\mathcal{C})$ is tractable.*

Any more general sufficient property for tractability would likely have to preserve this feature of making use of the width of the core without actually requiring the computation of the core. Hence, in light of Theorem 3 and Corollary 1 we conclude the section with the following conjecture.

Conjecture 1. *Let $\mathcal{C}$ be a class of non-exotic CSPs. Then $\mathrm{CSP}(\mathcal{C})$ is tractable if and only if $\mathcal{C}$ has bounded semantic hypertree width.*

5   Conclusion & Outlook

We have given characterizations of the fixed-parameter tractable classes of CSPs and UCQs. This allows us to determine the parameterized complexity of problems that have CSP or UCQ formulations by determining if the class of these formulations has bounded $\texttt{sem-}subw$. This motivates further work on theoretical tools that help to show whether a class has bounded $\texttt{sem-}subw$. We believe that further study of *adaptive width* [Marx, 2011], which is bounded iff $subw$ is bounded, can be a productive avenue of research here.

The characterization of polynomial time solvable CSPs remains open. We have motivated a new class of non-exotic problems that merits further research. In particular, we wish to resolve Conjecture 1, which we believe to be an important step towards the general problem.

Recent work has proposed the use of hybrid width parameters for the study of the computational complexity of CSP, e.g., [Ganian *et al.*, 2019]. Such hybrid width parameters, which consider both the query structure and database content, are a natural avenue for further research.

Moreover, we are intrigued by the connections to VC dimension, which is an important parameter in learnability theory [Blumer *et al.*, 1989]. We plan to further investigate the nature of the relationship between decomposition methods and learnability theory.

Acknowledgments

References

[Adler *et al.*, 2007] Isolde Adler, Georg Gottlob, and Martin Grohe. Hypertree width and related hypergraph invariants. *Eur. J. Comb.*, 28(8):2167–2181, 2007.

[Atserias *et al.*, 2006] Albert Atserias, Anuj Dawar, and Phokion G. Kolaitis. On preservation under homomorphisms and unions of conjunctive queries. *J. ACM*, 53(2):208–237, 2006.

[Barceló *et al.*, 2017] Pablo Barceló, Andreas Pieris, and Miguel Romero. Semantic optimization in tractable classes of conjunctive queries. *SIGMOD Record*, 46(2):5–17, 2017.

[Blumer *et al.*, 1989] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *J. ACM*, 36(4):929–965, 1989.

[Bulatov, 2017] Andrei A. Bulatov. A dichotomy theorem for nonuniform CSPs. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 319–330, 2017.

[Chen and Dalmau, 2005] Hubie Chen and Víctor Dalmau. Beyond hypertree width: Decomposition methods without decompositions. In *Principles and Practice of Constraint Programming - CP 2005, 11th International Conference, CP 2005, Sitges, Spain, October 1-5, 2005, Proceedings*, pages 167–181, 2005.

[Chen and Müller, 2015] Hubie Chen and Moritz Müller. The fine classification of conjunctive queries and parameterized logarithmic space. *TOCT*, 7(2):7:1–7:27, 2015.

[Chen, 2014] Hubie Chen. On the complexity of existential positive queries. *ACM Trans. Comput. Log.*, 15(1):9:1–9:20, 2014.

[Do and Kambhampati, 2001] Minh Binh Do and Subbarao Kambhampati. Planning as constraint satisfaction: Solving the planning graph by compiling it into CSP. *Artif. Intell.*, 132(2):151–182, 2001.

[Fagin *et al.*, 2005] Ronald Fagin, Phokion G. Kolaitis, and Lucian Popa. Data exchange: getting to the core. *ACM Trans. Database Syst.*, 30(1):174–210, 2005.

[Flum and Grohe, 2006] Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2006.

[Ganian *et al.*, 2019] Robert Ganian, Sebastian Ordyniak, and Stefan Szeider. A join-based hybrid parameter for constraint satisfaction. In *Principles and Practice of Constraint Programming - 25th International Conference, CP 2019, Stamford, CT, USA, September 30 - October 4, 2019, Proceedings*, pages 195–212, 2019.

[Gottlob *et al.*, 2000] Georg Gottlob, Nicola Leone, and Francesco Scarcello. A comparison of structural CSP decomposition methods. *Artif. Intell.*, 124(2):243–282, 2000.

[Gottlob *et al.*, 2002] Georg Gottlob, Nicola Leone, and Francesco Scarcello. Hypertree decompositions and tractable queries. *J. Comput. Syst. Sci.*, 64(3):579–627, 2002.

[Gottlob *et al.*, 2019] Georg Gottlob, Matthias Lanzinger, and Reinhard Pichler. Semantic width revisited (extended abstract). In *Proceedings of the 13th Alberto Mendelzon International Workshop on Foundations of Data Management, Asunción, Paraguay, June 3-7, 2019*, 2019.

[Gottlob *et al.*, 2020] Georg Gottlob, Matthias Lanzinger, Reinhard Pichler, and Igor Razgon. Complexity analysis of generalized and fractional hypertree decompositions. *CoRR*, abs/2002.05239, 2020.

[Grohe and Marx, 2014] Martin Grohe and Dániel Marx. Constraint solving via fractional edge covers. *ACM Trans. Algorithms*, 11(1):4:1–4:20, 2014.

[Grohe, 2007] Martin Grohe. The complexity of homomorphism and constraint satisfaction problems seen from the other side. *J. ACM*, 54(1):1:1–1:24, 2007.

[Hell and Nesetril, 1990] Pavol Hell and Jaroslav Nesetril. On the complexity of *H*-coloring. *J. Comb. Theory, Ser. B*, 48(1):92–110, 1990.

[Impagliazzo *et al.*, 2001] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.*, 63(4):512–530, 2001.

[Kolaitis and Vardi, 2000] Phokion G. Kolaitis and Moshe Y. Vardi. Conjunctive-query containment and constraint satisfaction. *J. Comput. Syst. Sci.*, 61(2):302–332, 2000.

[Kumar, 1992] Vipin Kumar. Algorithms for constraint-satisfaction problems: A survey. *AI Magazine*, 13(1):32–44, 1992.

[Maier, 1983] David Maier. *The Theory of Relational Databases*. Computer Science Press, 1983.

[Marx, 2011] Dániel Marx. Tractable structures for constraint satisfaction with truth tables. *Theory Comput. Syst.*, 48(3):444–464, 2011.

[Marx, 2013] Dániel Marx. Tractable hypergraph properties for constraint satisfaction and conjunctive queries. *J. ACM*, 60(6):42:1–42:51, 2013.

[Narváez, 2018] David E. Narváez. Constraint satisfaction techniques for combinatorial problems. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*, pages 8028–8029, 2018.

[Robertson and Seymour, 1986] Neil Robertson and Paul D. Seymour. Graph minors. II. algorithmic aspects of treewidth. *J. Algorithms*, 7(3):309–322, 1986.

[Sagiv and Yannakakis, 1980] Yehoshua Sagiv and Mihalis Yannakakis. Equivalences among relational expressions with the union and difference operators. *J. ACM*, 27(4):633–655, 1980.

[Zhuk, 2017] Dmitriy Zhuk. A proof of CSP dichotomy conjecture. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 331–342, 2017.