Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Monadic NIP in monotone classes of relational structures

(Extended abstract)

Samuel Braunfeld* Anuj Dawar† Ioannis Eleftheriadis‡
Aris Papadopoulos§

Abstract

We prove that for any monotone class of finite relational structures, the first-order theory of the class is NIP in the sense of stability theory if, and only if, the collection of Gaifman graphs of structures in this class is nowhere dense. This generalises results previously known for graphs to relational structures and answers an open question posed by Adler and Adler (2014). The result is established by the application of Ramsey-theoretic techniques and shows that the property of being NIP is highly robust for monotone classes. We also show that the model-checking problem for first-order logic is intractable on any monotone class of structures that is not (monadically) NIP. This is a contribution towards the conjecture that the hereditary classes of structures admitting fixed-parameter tractable model-checking are precisely those that are monadically NIP.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-029`

*Computer Science Institute of Charles University (IUUK), Praha, Czech Republic. E-mail: `sbraunfeld@iuuk.mff.cuni.cz`. Supported by Project 21-10775S of the Czech Science Foundation (GAČR), European Union's Horizon 2020 research and innovation programme (grant agreement No 810115 – Dynasnet).

†Department of Computer Science and Technology, University of Cambridge, UK. E-mail: `anuj.dawar@cl.cam.ac.uk`. Supported by EPSRC grant EP/T007257/1.

‡Department of Computer Science and Technology, University of Cambridge, UK. E-mail: `ie257@cam.ac.uk`. Supported by a George and Marie Vergottis Scholarship awarded through Cambridge Trust, an Onassis Foundation Scholarship, and a Robert Sansom Studentship.

§School of Mathematics, Univesity of Leeds, UK. E-mail: mmadp@leeds.ac.uk. Supported by a Leeds Doctoral Scholarship.

1 Introduction

The development of stability theory in classical model theory, originating with Shelah's classification programme fifty years ago [12, 2], has sought to distinguish *tame* first-order theories from *wild* ones. A key discovery is that combinatorial configurations serve as dividing lines in this classification.

Separately, in the development of finite model theory, there has been in interest in investigating *tame* classes of finite structures. Here tameness can refer to *algorithmic tameness*, meaning that algorithmic problems that are intractable in general may be tractable on a tame class; or it can refer to *model-theoretic tameness*, meaning that the class enjoys some desirable model-theoretic properties that are absent in the class of all finite structures. See [4] for an exposition of these notions of tameness. The tame classes that arise in this context are often based on notions taken from the study of sparse graphs [9] and usually extended to classes of relational structures beyond graphs by applying them to the *Gaifman graphs* of such structures.

In the context of algorithmic tameness of sparse classes, this line of work culminated in the major result of Grohe et al. [7] showing that the problem of model checking first-order sentences is fixed-parameter tractable (FPT) on any class of graphs that is *nowhere dense*. This generalized a sequence of earlier results showing the tractability of the model checking problem on classes of graphs satisfying other notions of sparsity. Moreover, it is also known [8] that this is the limit of tractability for *monotone* classes of graphs. That is to say that (under reasonable assumptions) any monotone class of graphs in which first-order model checking is FPT is necessarily nowhere dense. These results underline the centrality of the notion of nowhere denseness in the study of sparse graph classes.

A significant line of recent research has sought to generalize the methods and results on tame sparse classes of graphs to more general classes that are not necessarily sparse. Interestingly, this has tied together notions of tameness arising in finite model theory and those in classical model theory. Notions arising from stability theory play an increasingly important role in these considerations (see [10, 6], for example). Central to this connection is the realisation that for well-studied notions of sparseness in graphs, the first-order theory of a sparse class $\mathcal{C}$ is stable. Thus, stability-theoretic notions of tameness, applied to the theory of a class of finite structures, generalize the notions of tameness emerging from the theory of sparsity.

A key result connecting the two directions is that a monotone class of finite graphs is stable if, and only if, it is nowhere dense. This connection between stability and combinatorial sparsity was established in the context of infinite graphs by Podewski and Ziegler [11] and extended to classes of finite graphs by Adler and Adler [1]. Indeed, for monotone classes of graphs, stability is a rather robust concept as the theory of such a class is stable if, and only if, it is NIP, and these conditions on monotone classes are in turn equivalent to it being monadically stable and monadically NIP.

A question posed by Adler and Adler is whether their result can be extended from graphs to structures in any finite relational language. We settle this question in the present paper by establishing Theorem 1 below. In the following $\mathsf{Gaif}(\mathcal{C})$ denotes the collection of Gaifman graphs of structures in the class $\mathcal{C}$. Note that the extension from graphs to relational structures requires considerable combinatorial machinery in the form of Ramsey-theoretic results, which we detail in later sections. We also relate the characterization to the tractability of the classes. In summary, our key results are stated in the following theorem.

Theorem 1. *Let $\mathcal{C}$ be a monotone class of finite structures in a finite relational language. Then, the following are equivalent:*

1. *$\mathcal{C}$ is NIP;*
2. *$\mathcal{C}$ is monadically NIP;*
3. *$\mathcal{C}$ is stable;*
4. *$\mathcal{C}$ is monadically stable;*
5. $\mathsf{Gaif}(\mathcal{C})$ *is nowhere dense; and*
6. *(assuming $\mathsf{AW}[*] \neq \mathsf{FPT}$) $\mathcal{C}$ admits fixed-parameter tractable model checking.*

Moreover, the equivalence of the first six notions also holds for classes containing infinite structures.

The equivalence of the first four notions for any monotone class $\mathcal{C}$ is due to Braunfeld and Laskowski [3]. The equivalence of the fifth and sixth notions follows by results in sparsity theory (see [9]). We, therefore, establish the equivalence of the first with the fifth and the sixth. More precisely, we show that if $\mathsf{Gaif}(\mathcal{C})$ is not nowhere dense, then $\mathcal{C}$ admits a formula with the independence property. That nowhere density of $\mathsf{Gaif}(\mathcal{C})$ implies tractability is implicit in [7]. We establish the converse of this statement here.

2 Preliminaries

We assume familiarity with first-order logic and the basic concepts of model theory and graph theory. Throughout this paper, $\mathcal{L}$ denotes a finite, first-order, relational language. Given an $\mathcal{L}$-structure M, we write $\mathsf{Gaif}(M)$ for the *Gaifman graph* of M, i.e. the graph on domain M with the property that two elements are adjacent if and only if they appear together in a relation of M. For a class $\mathcal{C}$ of $\mathcal{L}$-structures, we write $\mathsf{Gaif}(\mathcal{C})$ for the class of graphs $\{\mathsf{Gaif}(M) : M \in \mathcal{C}\}$. We say that a class $\mathcal{C}$ is monotone if it is closed under weak substructures, i.e. if $(M, R^M)_{R\in\mathcal{L}} \in \mathcal{C}$ then $(M', R^{M'})_{R\in\mathcal{L}} \in \mathcal{C}$ for any $M' \subseteq M$ and $R^{M'} \subseteq R^M$.

We say that a class $\mathcal{C}$ of $\mathcal{L}$-structures is *NIP* (Not the Independence Property) if there is no $\mathcal{L}$-formula $\phi(\bar{x}, \bar{y})$ satisfying that for all bipartite graphs $G = (U, V; E) \in \mathfrak{B}$ there is some $M_G \in \mathcal{C}$ and sequences of tuples $(\bar{a}_i)_{i\in U}$ and $(\bar{b}_j)_{j\in V}$ such that:

$$M_G \models \phi(\bar{a}_i, \bar{b}_j) \text{ if, and only if, } (i, j) \in E.$$

Moreover, a class $\mathcal{C}$ of graphs is said to be *nowhere dense* if for every $r \in \mathbb{N}$ there is some $n \in \mathbb{N}$ such that for all $G \in \mathcal{C}$ we have that $K_n^{(r)}$ is not a subgraph of G, and otherwise, $\mathcal{C}$ is *somewhere dense.*

By the model checking problem on a class $\mathcal{C}$, we refer to the parametrised decision problem whereby, given a structure $M \in \mathcal{C}$ and an FO-sentence ϕ whose depth acts as parameter, we want to decide if M satisfies ϕ. We say that the model checking problem on a class $\mathcal{C}$ is *fixed-parameter tractable*, if there is an algorithm that decides on input (M, ϕ) whether $M \models \phi$, in time $f(|\phi|) \cdot |M|^{\mathcal{O}(1)}$ for some computable function f. Model checking on the class of all graphs is complete with respect to the complexity class $\mathsf{AW}[*]$, which is conjectured to strictly contain the class FPT. We shall assume throughout that $\mathsf{AW}[*] \neq \mathsf{FPT}$.

3 Main results

Here, we sketch the proofs of implications (1) $\implies$ (5) and (6) $\implies$ (5) from Theorem 1. We first prove that for any monotone class $\mathcal{C}$ of relational structures whose Gaifman class is somewhere dense, there is a formula which codes the edge relation of all bipartite graphs uniformly over $\mathcal{C}$. We work towards this theorem via a preparatory lemma, which has the benefit of applying to classes that are not necessarily monotone. Intuitively, this tells us that in any class of relational structures $\mathcal{C}$ whose Gaifman class is somewhere dense, there is a primitive positive formula which codes the edge relation of any complete bipartite graph with "sufficiently disjoint" witnesses.

Lemma 1. *Let $\mathcal{C}$ be a class of $\mathcal{L}$-structures such that $\mathsf{Gaif}(\mathcal{C})$ is somewhere dense. Then there is a primitive positive formula $\phi(\bar{x}, \bar{y}, \bar{z}) = \exists \bar{w} \psi(\bar{x}, \bar{y}, \bar{z}, \bar{w})$ with parameters $\bar{p}$, and for every $n \in \mathbb{N}$ there is some $M_n \in \mathcal{C}$ and tuples $(\bar{a}_i)_{i \in [n]}$, $(\bar{b}_j)_{j \in [n]}$, $(\bar{c}_{i,j})_{(i,j) \in [n]^2}$, $(\bar{d}_{i,j})_{i,j \in [n]^2}$ from M_n such that the following hold for all $i, i', j, j' \in [n]$:*

1. *$M_n \models \psi(\bar{a}_i, \bar{b}_j, \bar{c}_{i,j}, \bar{d}_{i,j})$;*
2. *$\bar{a}_i(k) \neq \bar{a}_{i'}(k)$, for $i \neq i'$ and all $k \in [|\bar{x}|]$;*
3. *$\bar{b}_j(k) \neq \bar{b}_{j'}(k)$, for $j \neq j'$ and all $k \in [|\bar{y}|]$;*
4. *$\bar{c}_{i,j}(k) \neq \bar{c}_{i',j'}(k)$ and $\bar{c}_{i,j}(k) \neq \bar{c}_{i,j}(l)$, for $(i,j) \neq (i',j')$ and all $k \neq l$ from $[|\bar{z}|]$;*
5. *$\bar{d}_{i,j}(k) \neq \bar{d}_{i',j'}(k)$, for $(i,j) \neq (i',j')$ and all $k \in [|\bar{w}|]$.*

The proof of this lemma is a combinatorial argument, resting on few applications of different Ramsey theorems. First, we ensure that the subdivided edges coming from the assumption that $\mathsf{Gaif}(\mathcal{C})$ is somewhere dense are witnessed in every M_n by the same sequence of relations $R_1, \ldots, R_k$; this requires an application of the finite Ramsey theorem (for local uniformity within each structure), and of the pigeonhole principle (for global uniformity for the whole class). Next, by consecutive applications of the canonical Erdős-Rado theorem (see [5]), we may obtain a finite set of elements that are common in all such relations, and ensure that the remaining elements on these are essentially distinct. The level of disjointedness achieved is precisely what allows us, under the additional assumption that $\mathcal{C}$ is monotone, to remove relations so as to turn the encoded complete bipartite graphs into arbitrary bipartite graphs and violate NIP. Consequently, we establish the following.

Theorem 2. *Let $\mathcal{C}$ be a monotone class of $\mathcal{L}$-structures such that $\mathsf{Gaif}(\mathcal{C})$ is somewhere dense. Then there is a primitive positive formula $\phi(\bar{x}, \bar{y}) = \exists \bar{w} \psi(\bar{x}, \bar{y}, \bar{w})$ with parameters $\bar{p}$ and for each bipartite graph $G = (U, V; E)$ there is some $M_G \in \mathcal{C}$ and sequences of tuples $(\bar{a}_u)_{u \in U}$ $(\bar{b}_v)_{v \in V}$, $(\bar{h}_{u,v})_{(u,v) \in E}$ from M_G such that:*

1. *$M_G \models \phi(\bar{a}_u, \bar{b}_v)$ if, and only if, $(u, v) \in E$ (so, in particular $\mathcal{C}$ is not NIP);*
2. *If $(u, v) \in E$ then $M_G \models \psi(\bar{a}_u, \bar{b}_v, \bar{h}_{u,v})$;*
3. *The equality type of $\bar{p}_{u,v} = \bar{a}_u {}^\frown \bar{b}_v {}^\frown \bar{h}_{u,v}$ is constant for all $(u, v) \in E(G)$;*
4. *Any two tuples in $\{\bar{a}_u, \bar{b}_v, \bar{h}_{u,v} : u \in U, v \in V\}$ are disjoint and do not intersect the parameters $\bar{p}$.*

Next, we prove that any monotone class of relational structures whose Gaifman class is somewhere dense polynomially interprets the class of all bipartite graphs, and is therefore intractable. Towards this, we first strengthen Theorem 2 to obtain a “simple path formula" that performs the encoding; this is essentially a primitive positive formula $\phi(\bar{x}, \bar{y})$ that describes a sequence of relation $R_1, \dots, R_k$, with the property that $\bar{x} \subseteq R_1$ and $\bar{y} \subseteq R_k$. Moreover, having full control over the equality type of the elements in M_G allows to obtain a polynomial-time construction of M_G from G.

Lemma 2. *Let $\mathcal{C}$ be a monotone class of $\mathcal{L}$-structures such that $\mathsf{Gaif}(\mathcal{C})$ is somewhere dense. Then there is a simple path formula $\phi(\bar{x}, \bar{y})$ with parameters $\bar{p}$ and a polynomial time computable function $\Phi : \mathfrak{B} \to \mathcal{C}$, such that for each bipartite graph $G = (U, V; E) \in \mathfrak{B}$ there are tuples $(\bar{a}_u)_{u \in U}$ $(\bar{b}_v)_{v \in V}$, $(\bar{h}_{u,v})_{(u,v) \in E}$ from $\Phi(G)$ satisfying:*

$$\Phi(G) \models \phi(\bar{a}_u, \bar{b}_v) \text{ if, and only if, } (u, v) \in E.$$

With this, we proceed to show intractability for monotone classes with somewhere Gaifman class. Our proof is essentially based on the proof of [8, Theorem 6.1], which covers the case of graphs. There, the aim is to definably distinguish the native points of an r-subdivided graph G from the subdivision points. The idea is to distinguish points by their degrees; however, while all subdivision points have degree two, other points in G may as well have degree two. To address this, we first pre-process G to obtain a graph G' by adding two pendant vertices to each non-isolated vertex. Then, G is definably recovered from G', and moreover, given an r-subdivision of G', we can definably distinguish the subdivision points and the remaining points by their degrees. Our construction is essentially the same, although the degree of a subdivision point is bounded by the length of paths in the subdivision, rather than by two.

Theorem 3. *Let $\mathcal{C}$ be a monotone class of $\mathcal{L}$-structures such that $\mathsf{Gaif}(\mathcal{C})$ is somewhere dense, and assume that $\mathsf{AW}[*] \neq \mathsf{FPT}$. Then FO model-checking on $\mathcal{C}$ is not fixed-parameter tractable.*

References

[1] Hans Adler and Isolde Adler. Interpreting nowhere dense graph classes as a classical notion of model theory. *European Journal of Combinatorics*, 36:322–330, 2014.

[2] John T. Baldwin. *Fundamentals of Stability Theory.* Perspectives in Logic. Cambridge University Press, 2017.

[3] Samuel Braunfeld and Michael C. Laskowski. Existential characterizations of monadic NIP, 2022.

[4] Anuj Dawar. Finite model theory on tame classes of structures. In *MFCS*, volume 4708 of *Lecture Notes in Computer Science*, pages 2–12. Springer, 2007.

[5] Paul Erdős and Richard Rado. A combinatorial theorem. *Journal of the London Mathematical Society*, 1(4):249–255, 1950.

[6] Jakub Gajarský, Michal Pilipczuk, and Szymon Torunczyk. Stable graphs of bounded twin-width. In Christel Baier and Dana Fisman, editors, *LICS '22: 37th Annual ACM/IEEE Symposium on Logic in Computer Science, Haifa, Israel, August 2 - 5, 2022*, pages 39:1–39:12. ACM, 2022.

[7] Martin Grohe, Stephan Kreutzer, and Sebastian Siebertz. Deciding first-order properties of nowhere dense graphs. *J. ACM*, 64(3), jun 2017.

[8] Stephan Kreutzer and Anuj Dawar. Parameterized complexity of first-order logic. *Electron. Colloquium Comput. Complex.*, TR09-131, 2009.

[9] Jaroslav Nešetřil and Patrice Ossona De Mendez. *Sparsity: graphs, structures, and algorithms*, volume 28. Springer Science & Business Media, 2012.

[10] Michal Pilipczuk, Sebastian Siebertz, and Szymon Torunczyk. Parameterized circuit complexity of model-checking on sparse structures. In Anuj Dawar and Erich Grädel, editors, *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018*, pages 789–798. ACM, 2018.

[11] Klaus-Peter Podewski and Martin Ziegler. Stable graphs. *Fund. Math*, 100(2):101–107, 1978.

[12] Saharon Shelah. *Classification theory and the number of nonisomorphic models*, volume 92 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam-New York, 1978.