Article:

Adler, H and Adler, I (2014) Interpreting nowhere dense graph classes as a classical notion of model theory. European Journal of Combinatorics, 36. pp. 322-330. ISSN 0195-6698

https://doi.org/10.1016/j.ejc.2013.06.048

Interpreting nowhere dense graph classes as a classical notion of model theory

Hans Adler
Vienna University, Austria
hans.adler@univie.ac.at

Isolde Adler
Goethe University Frankfurt, Germany
iadler@informatik.uni-frankfurt.de

June 18, 2013

Abstract

A class of graphs is nowhere dense if for every integer r there is a finite upper bound on the size of complete graphs that occur as r-minors. We observe that this recent tameness notion from (algorithmic) graph theory is essentially the earlier stability theoretic notion of superflatness. For subgraph-closed classes of graphs we prove equivalence to stability and to not having the independence property. Expressed in terms of PAC learning, the concept classes definable in first-order logic in a subgraph-closed graph class have bounded sample complexity, if and only if the class is nowhere dense.

1 Introduction

Recently, Nešetřil and Ossona de Mendez [14, 13] introduced *nowhere dense* classes of finite graphs, a generalisation of many natural and important classes such as graphs of bounded degree, planar graphs, graphs excluding a fixed minor and graphs of bounded expansion. These graph classes play an important role in algorithmic graph theory, as many computational problems that are hard in general become tractable when restricted to such classes. All these graph classes are nowhere dense. Dawar and Kreutzer [4] gave efficient algorithms for domination problems on classes of nowhere dense graphs. Moreover, nowhere dense classes were studied earlier in the area of finite model theory [2, 3] under the guise of *(uniformly) quasi-wide* classes and also turned out to be well-behaved.[1] The book [11] covers the recent results in this area.

[1] The equivalence of nowhere dense, quasi-wide and uniformly quasi-wide for subgraph-closed classes was proved by Nešetřil and Ossona de Mendez [13].

In this paper we observe that nowhere density is essentially the stability theoretic notion of superflatness which was introduced by Podewski and Ziegler [12] in 1978 because of its connection to stability.

For some time we have been looking for a way to translate between tameness in finite model theory and in stability theory. A key obstacle was the fact that tameness notions in finite model theory are generally not even invariant under taking the complement of a relation, whereas in stability theory the exact choice of signature does not matter and all first-order definable sets are a priori equal. It now appears that on graph classes such that every (not necessarily induced) subgraph of a member is again in the class, tameness notions from stability theory, finite model theory and also algorithmic graph theory can be compared in a meaningful though somewhat coarse way. For subgraph-closed classes of graphs we show that nowhere density is equivalent to stability and to dependence (not having the independence property, or, equivalently, all first-order formulas having finite Vapnik-Chervonenkis dimension).

The equivalence of stability and dependence in this context is somewhat surprising, although it is well known under the stability theoretic assumption of simplicity. Stability and the independence property are two key dividing lines in Shelah's classification theory programme for infinite model theory [15, 6]. A theory is stable if it does not have a first-order formula that codes an infinite linear order on a set of tuples. This strong and robust tameness property is the key assumption on which Shelah originally built his monumental machinery of stability theory. At the other end, theories with the independence property have a formula that can code every subset of some infinite set.[2] Stability theory has recently made advances into general theories without the independence property (see last section of [1] for pointers), although much remains to be done. The independence property is a strong wildness property, even though some theories with the independence property, such as that of the random graph, are actually very easy to understand from a stability-theoretic point of view. A formula has the independence property if and only if it has infinite Vapnik-Chervonenkis dimension – a key wildness notion in computational learning theory [16, 10].

We hope for further translations between notions of tameness in stability theory and notions of tameness in combinatorial graph theory. This should allow us to identify well-behaved graph classes with good algorithmic properties. Moreover, we hope that these translations can ultimately be refined and extended to more general contexts such as

[2] Since we will apply the notions of stability and independence to graphs, we will be careful not to use their unrelated graph theoretic homonyms.

arbitrary classes of relational structures.

2 Shallow graph minors and nowhere density

In this paper graphs are undirected, without loops or multiple edges, and not necessarily finite. Notationally, we do not distinguish between a graph G and its vertex set (otherwise often denoted by $V(G)$). From the point of view of model theory, a graph is a relational structure G with an irreflexive and symmetric binary relation $E^G \subseteq G^2$. Thus each edge $\{a, b\}$ of G is represented by two directed edges $(a, b), (b, a) \in E^G$. For the standard notions of graph theory we refer the reader to Diestel's book [5].

H is a *minor* of G if there is a subgraph $U \subseteq G$ (not necessarily an induced subgraph) and an equivalence relation ϵ on U with connected classes, such that $H \cong U/\epsilon$, i.e. H is the result of contracting each ϵ-class to a single vertex in such a way that two distinct ϵ-classes are connected by an edge if and only if there exist two respective members that are connected by an edge. H is an *r-minor* of G if each ϵ-equivalence class contains a vertex from which the other vertices have distance at most r.

H is a *topological minor* of G if there is a subgraph $U \subseteq G$ (not necessarily an induced subgraph) and an equivalence relation ϵ on E^U, i.e. on the edges, such that each ϵ-class is a path whose interior vertices all have degree 2, and $H \cong U/\epsilon$, i.e. H is the result of contracting each ϵ-class to a single edge. H is a *topological r-minor* of G if moreover each ϵ-equivalence class consists of at most $2r+1$ edges. In other words, up to isomorphism the vertices of a topological (r-)minor H of G form a subset of the vertices of G (the branch vertices), and the edges of H correspond to pairwise internally vertex disjoint paths in G (of length at most $2r + 1$), whose interior points (subdividing vertices) avoid H.

In the following, we will consider isomorphism-closed classes $\mathcal{C}$ of graphs. A (topological) (r-)minor of $\mathcal{C}$ is a (topological) (r-)minor of a member of $\mathcal{C}$, respectively. We write $\mathcal{C} \,\nabla\, r$ for the class of r-minors of $\mathcal{C}$, and $\mathcal{C} \,\widetilde{\nabla}\, r$ for the class of topological r-minors of $\mathcal{C}$. In particular, $\mathcal{C} \,\nabla\, 0 = \mathcal{C} \,\widetilde{\nabla}\, 0$ is the class of all graphs isomorphic to a subgraph of a member of $\mathcal{C}$. Also note $\mathcal{C} \,\widetilde{\nabla}\, r \subseteq \mathcal{C} \,\nabla\, r$, $(\mathcal{C} \,\nabla\, r) \,\nabla\, s \subseteq \mathcal{C} \,\nabla (2rs + r + s)$ and $(\mathcal{C} \,\widetilde{\nabla}\, r) \,\widetilde{\nabla}\, s \subseteq \mathcal{C} \,\widetilde{\nabla} (2rs + r + s)$.

Nešetřil and Ossona de Mendez proved that as r goes to infinity, there are only three possible asymptotic behaviours for the growth of the number of edges of finite r-minors, or equivalently finite topological r-minors, in terms of their number of vertices: finitely bounded, linear, or quadratic.

Fact 1.

$$\lim_{r\to\infty} \limsup_{\substack{H\in\mathcal{C}\,\nabla\, r\\ H\text{finite}\\ |H|\to\infty}} \frac{\log\left|E^H\right|}{\log|H|} \quad = \quad \lim_{r\to\infty} \limsup_{\substack{H\in\mathcal{C}\,\widetilde{\nabla}\, r\\ H\text{finite}\\ |H|\to\infty}} \frac{\log\left|E^H\right|}{\log|H|} \quad \in \quad \{0,1,2\},$$

where $\left|E^H\right|$ *and* $|H|$ *are the number of edges and the number of vertices of* H*, respectively. Moreover, the quadratic case (right-hand side 2) is equivalent to the statement that for some* r *there is no finite upper bound on the sizes of complete graphs that occur as* r*-minors of* $\mathcal{C}$*, or equivalently as topological* r*-minors.*

They called $\mathcal{C}$ *nowhere dense* in the linearly bounded case, i.e. when for every r there is a finite upper bound for the sizes of complete graphs that occur as r-minors (or, equivalently, topological r-minors) of $\mathcal{C}$ [14]. If $\mathcal{C}$ is nowhere dense, then so is every subclass of every class of the form $\mathcal{C}\,\nabla\, r$.[3]

An *m-clique* is a complete graph on m vertices, denoted by K_m. By K_m^r we denote the result of subdividing each edge of the m-clique K_m exactly r times. Essentially following Podewski and Ziegler [12], we call $\mathcal{C}$ *superflat* if for every r there is an m such that K_m^r does not occur as a subgraph of a member of $\mathcal{C}$. Using the finite Ramsey theorem (see e.g. [5]), it is easy to see:

Remark 2. Let $\mathcal{C}$ be a class of graphs. $\mathcal{C}$ is nowhere dense if and only if $\mathcal{C}$ is superflat.

3 Stability of graphs

Graphs and digraphs are examples of relational structures in the sense of first-order logic and model theory. Since we will later treat coloured digraphs in this framework, it is worth introducing some of the terminology in its general form. A *relational signature* is a set σ of relation symbols. Every relation symbol $R\in\sigma$ has an associated non-negative integer $\mathrm{ar}(R)$, its *arity.* A σ*-structure* M consists of a non-empty set U called the *universe* or *underlying set*, and a relation $R^M\subseteq U^{\mathrm{ar}(R)}$ for every $R\in\sigma$.[4] It is customary to write $a\in M$ instead of $a\in U$. For tuples $\bar{a}\in U^{|\bar{a}|}$, we simply write $\bar{a}\in M$ (slightly abusing notation). For the standard notions of model theory see the book by Hodges [9]. An

[3]Nešetřil and Ossona de Mendez only consider classes of finite graphs, but the version presented here is an obvious generalisation.

[4]Nullary relation symbols act syntactically and semantically like the variables of propositional logic, encoding Boolean variables within structures. Every nullary relation is a subset of the 1-element set which has the 0-tuple as its only element. Such a relation is true if and only if it is non-empty. Some authors exclude nullary relation symbols from the definition, but they may turn out useful in our context. It is safe to ignore this footnote.

undirected graph is an $\{E\}$-structure G with a single, binary relation E^G that is irreflexive and symmetric.

The formulas of *first-order logic* are built in the usual way from variables $x, y, z, x_1, \ldots$, the equality symbol $=$, the relation symbols in σ, the Boolean connectives $\wedge, \vee, \neg, \rightarrow$, and the quantifiers $\forall, \exists$ ranging over the universe of the structure. A *free* variable in a first-order formula is a variable x that occurs outside the scope of a quantifier $\forall x$ or $\exists x$. The notation $\varphi(\bar{x})$ or $\varphi(x_1, \ldots, x_n)$ indicates that all free variables of the formula φ occur in $\bar{x} = (x_1, \ldots, x_n)$, a tuple without repetitions. For a formula $\varphi(x_1, \ldots, x_n)$, a structure M and elements $a_1, \ldots, a_n$ of the universe of M we write $M \models \varphi(a_1, \ldots, a_n)$ to say that M satisfies φ if the variables $x_1, \ldots, x_n$ are interpreted by the elements $a_1, \ldots, a_n$, respectively. Moreover, we let $\varphi(\bar{x})^M := \{\bar{a} \in M \mid M \models \varphi(\bar{a})\}$.

For example, we can use the following first-order formula to recognise a dominating set of size k in a graph. A set of vertices $X \subseteq G$ is called a *dominating set* in G, if every vertex of G is either in X or has a neighbour in X. Let

$$\varphi_{\mathrm{DS}}(x_1, \ldots, x_k) := \Big(\bigwedge_{1 \le i < j \le k} \neg x_i = x_j \Big) \wedge \forall y \Big(\bigvee_{i=1}^{k} y = x_i \vee \bigvee_{i=1}^{k} E y x_i \Big).$$

Then for any graph G and vertices $a_1, \ldots, a_k \in G$ we have $G \models \varphi_{\mathrm{DS}}(a_1, \ldots, a_k)$ if and only if $\{a_1, \ldots, a_k\}$ is a dominating set of size k in G.

Let $\mathcal{C}$ be a class of structures of a fixed signature. A first-order formula $\varphi(\bar{x}, \bar{y})$ is said to have the *order property* with respect to $\mathcal{C}$ if it has the *n-order property* for all n, i.e. if for every n there exist a structure $M \in \mathcal{C}$ and tuples $\bar{a}_0, \ldots, \bar{a}_{n-1}, \bar{b}_0, \ldots, \bar{b}_{n-1} \in M$ such that $M \models \varphi(\bar{a}_i, \bar{b}_j)$ holds if and only if $i < j$. A class $\mathcal{C}$ of structures is called *stable* if there is no such formula with respect to $\mathcal{C}$. It is easy to see that $\mathcal{C}$ is stable if and only if there is no formula $\psi(\bar{u}, \bar{v})$ with $|\bar{u}| = |\bar{v}|$, such that for every n there exist a structure $M \in \mathcal{C}$ and tuples $\bar{c}_0, \ldots, \bar{c}_{n-1} \in M$ such that $M \models \psi(\bar{c}_i, \bar{c}_j)$ holds if and only if $i \le j$, i.e. ψ orders the tuples linearly.

Stability and the (n-)order property come from stability theory [15, 6], where they are defined for the class of models of a complete first-order theory. A single structure M is called stable if $\{M\}$ is stable. This is equivalent to requiring that the class of all structures elementarily equivalent to M be stable, and so our notion of stability generalises the usual one. In this paper we are primarily interested in applying the concept to classes of finite graphs.

In this paper, we need to work with first-order interpretations with parameters.[5] Fortunately, we will get away with a much simpler (yet

[5] In the most general case there can be a tuple of parameters instead of the single

still complicated) special case.

For our purposes, a *first-order interpretation* $I\colon \mathcal{C} \to \mathcal{D}$ of a class $\mathcal{C}$ of σ-structures in a class $\mathcal{D}$ of τ-structures consists of the following data:

- for each $A \in \mathcal{C}$ a structure $M_I^A \in \mathcal{D}$,
- for each $A \in \mathcal{C}$ an element $c_I^A \in M_I^A$,
- for each structure $A \in \mathcal{C}$ a map $f_I^A\colon A \to M_I^A$,
- a τ-formula $\delta_I(x, z)$,
- for each σ-formula $\varphi(\bar{x})$ a τ-formula $\varphi_I(\bar{x}, z)$.

Moreover, for an interpretation it is required that

- every $A \in \mathcal{C}$ satisfies $f_I^A(A) = \delta_I(x, c_I^A)^{M_I^A}$, and
- for any $A \in \mathcal{C}$, σ-formula $\varphi(\bar{x})$ and compatible tuple $\bar{a} \in A$,

$$A \models \varphi(\bar{a}) \iff M_I^A \models \varphi_I(f_I^A(\bar{a}), c_I^A).$$

It is sufficient to have $\varphi_I(\bar{x}, z)$ given just for formulas $\varphi(\bar{x})$ without quantifiers or connectives. Formulas φ_I for more general φ can then be constructed by induction on the structure (or length) of a formula.

Remark 3. If $\mathcal{C}$ is interpretable in $\mathcal{D}$ and $\mathcal{D}$ is stable, then so is $\mathcal{C}$.

The notion of superflatness was originally introduced by Podewski and Ziegler as a simple sufficient condition for stability of infinite graphs. A graph G is *superflat* in their sense if and only if $\{G\}$ is superflat.

Fact 4 (Podewski, Ziegler [12])**.** *Every superflat graph G is stable.*

Every subgraph of a superflat graph is superflat, but every graph is a subgraph of a complete graph, and complete graphs are stable in the sense of model theory. Therefore the converse of Fact 4 does not hold.

Lemma 5. *Let $\mathcal{C}$ be a class of graphs. If $\mathcal{C}$ is superflat, then $\mathcal{C}$ is stable.*

Proof. The basic idea is as follows. Let $G(\mathcal{C}) = \dot{\bigcup}_{A\in\mathcal{C}} A$ denote the graph which is the disjoint union of all graphs in $\mathcal{C}$. If $\mathcal{C}$ is superflat, then so is $G(\mathcal{C})$, and by Fact 4 the graph $G(\mathcal{C})$ is stable.

If all graphs in $\mathcal{C}$ are connected and have a uniform upper bound on the diameter, then it is not hard to describe an interpretation of $\mathcal{C}$ in $\{G(\mathcal{C})\}$, and to conclude that $\mathcal{C}$ is stable.

---

parameter c_I^A, and the maps f^A go to $(M_I^A)^n/\varepsilon$, for some $n \in \mathbb{N}$ and an equivalence relation ε which is definable by a formula with parameters.

With no such connectedness assumptions, we have to add another step. From every graph $A \in \mathcal{C}$ we derive a graph A' containing A as an induced subgraph. A' has one additional vertex c^A, and c^A has an edge to every vertex of A. Let $\mathcal{C}' := \{A' \mid A \in \mathcal{C}\}$. It is easy to see that $\mathcal{C}'$ and $G(\mathcal{C}')$ are again superflat. Hence $G(\mathcal{C}')$ is stable by Fact 4.

The following describes an interpretation $I\colon \mathcal{C} \to \{G(\mathcal{C}')\}$.

- $M_I^A = G(\mathcal{C}')$ for all $A \in \mathcal{C}$,
- f_I^A is the obvious map from A to its copy in $G(\mathcal{C}')$,
- $c_I^A = c^A$,
- $\delta_I(x, z)$ is the formula $E(x, z)$,
- $\varphi_I(\bar{x}, z)$ is the relativisation of $\varphi(\bar{x})$ to the set of neighbours of z. E. g. if $\varphi(x_1, x_2, x_3) = \forall x_1 \exists x_2 \exists x_3 \big(E(x_1, x_2) \wedge E(x_2, x_3) \wedge E(x_1, x_3)\big)$, then the formula $\varphi_I(x_1, x_2, x_3, z)$ can be taken to be

$$\forall x_1 \big(E(x_1, z) \rightarrow \exists x_2 \big(E(x_2, z) \wedge \exists x_3 \big(E(x_3, z) \wedge \\ [E(x_1, x_2) \wedge E(x_2, x_3) \wedge E(x_1, x_3)]\big)\big)\big).$$

It is easy to see that $f_I^A(A) = \delta_I(x, c^A)^{G(\mathcal{C}')}$ and that $A \models \varphi(\bar{a}) \iff G(\mathcal{C}') \models \varphi_I(\bar{a}, c^A)$ for all $\bar{a} \in A$, so that I is in fact an interpretation. As $\mathcal{C}$ is interpretable in $\{G(\mathcal{C})\}$ and $G(\mathcal{C})$ is stable, by Remark 3 the class $\mathcal{C}$ is also stable. $\square$

4 Stability of coloured digraphs

In this section we extend Lemma 5 to classes of vertex- and edge-coloured directed graphs. More precisely, we extend it to relational structures where all relation symbols are at most binary.

By a *coloured digraph* we will understand a relational structure whose relation symbols are at most binary. The *underlying graph* or *Gaifman graph* $\underline{M}$ of a relational structure M is the graph with vertices the elements of M and edges all pairs $\{a, b, \}$ such that $a \neq b$ and a and b appear together in an instance of a relation of M. (I.e. $\{a, b\}$ is an edge of $\underline{M}$ if and only if $a \neq b$ and there exist a relation symbol R and a tuple $(c_1, \ldots, c_n) \in R^M$ such that $a = c_i$, $b = c_j$ for some i, j.) For every class $\mathcal{C}$ of structures M we let $\underline{\mathcal{C}}$ be the class of underlying graphs $\underline{M}$. For a class of structures $\mathcal{C}$, the combinatorial complexity of the graphs in $\underline{\mathcal{C}}$ is a good indication for the computational complexity of algorithmic problems on $\mathcal{C}$. This has been exploited in various areas such as complexity theory, database theory, constraint satisfaction, algorithmic graph theory and finite model theory. Here we will only consider underlying graphs of coloured digraphs, in which case the construction amounts to forgetting the colours, loops, and edge directions.

Lemma 6. *Every class $\mathcal{C}$ of coloured digraphs of a fixed countable signature can be interpreted in a class $\mathcal{C}'$ of undirected graphs such that $\mathcal{C}'$ is superflat if and only if $\underline{\mathcal{C}}$ is superflat.*

Proof. We enumerate the binary relation symbols in the signature of $\mathcal{C}$ as R_1, R_2, R_3, …, the unary relation symbols as P_1, P_2, P_3, … and the nullary relation symbols as A_1, A_2, A_3, …. For a single coloured digraph $G \in \mathcal{C}$, we define the following graph G'.

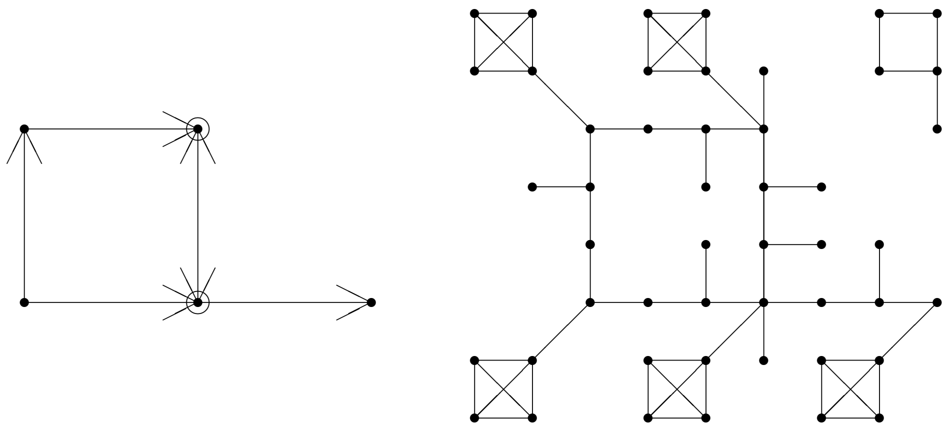

Figure 1: Proof of Lemma 6. Example of G with one unary and one binary relation (depicted by circles and arrows, respectively), and corresponding G'. The small connected component in the top right corner consists of the vertices d, d_0, d_1, d_2, d_3 and is recognisable as the only chordless 4-cycle in G'. As there are no nullary relation symbols in the signature, it serves no real purpose in this particular example.

All vertices of G are also vertices of G'. Moreover, for every vertex a of G there are four new vertices a_0, a_1, a_2, a_3 of G' and edges (a, a_0), (a_0, a_1), (a_0, a_2), (a_0, a_3), (a_1, a_2), (a_1, a_3), $(a_2, a_3) \in E^{G'}$. (To make the graph undirected, we also need the opposite edges (a_1, a_0), (a_2, a_0), (a_3, a_0), (a_2, a_1), (a_3, a_1), $(a_3, a_2) \in E^{G'}$. From now on we will not make this explicit.) These 4-cliques will allow us to pick out the vertices of G inside G' by means of a first-order formula, since there will be no other 4-cliques in G' and the 4-clique vertices will have no further connections. For every vertex $a \in P_i^G$ we add to G' new vertices $c_1, \ldots, c_i$ and the path (a, c_1), (c_1, c_2), …, $(c_{i-1}, c_i) \in E^{G'}$.

For every unordered pair $\{a, b\}$ of vertices from G (we allow $a = b$) such that (a, b) or (b, a) appears in one of the binary relations of G, G' contains new vertices c_{ba} and c_{ab} as well as edges (a, c_{ba}), (c_{ba}, c_{ab}), $(c_{ab}, b) \in E^{G'}$. I.e. any two vertices between which there is a directed edge are connected by an undirected path of length 3. For every directed edge $(a, b) \in R_i^G$ we add to G' new vertices $c_1, \ldots, c_i$ and the path $(c_{ab}, c_1), (c_1, c_2), \ldots, (c_{i-1}, c_i) \in E^{G'}$.

Finally, to treat the nullary relations in a similar way to the unary and binary cases, we add new vertices d, d_0, d_1, d_2, d_3 and edges

$$(d, d_0), (d_0, d_1), (d_1, d_2), (d_2, d_3), (d_3, d_0) \in E^{G'}.$$

For each i such that $G \models A_i$ we attach a new path $c_1, c_2, \ldots, c_i$ to d.

We get $\mathcal{C}'$ from $\mathcal{C}$ by treating each $G \in \mathcal{C}$ in this way. It is easy to check that $\mathcal{C}$ can be interpreted in $\mathcal{C}'$ and that $\mathcal{C}'$ is superflat if and only if $\underline{\mathcal{C}}$ is superflat. (For the interpretation, we do not actually need c_I^A in this case, so we can use formulas $\varphi_I(\bar{x}, z)$ in which the variable z does not actually occur, and choose arbitrary $c_I^A \in A$.) $\square$

Theorem 7. *Let $\mathcal{C}$ be a class of coloured digraphs of a fixed signature. If $\underline{\mathcal{C}}$ is superflat, then $\mathcal{C}$ is stable.*

Proof. Since every formula contains only a finite part of the signature, we may assume that the signature is finite. By Lemma 6 we can interpret $\mathcal{C}$ in a superflat class $\mathcal{C}'$ of graphs. By Lemma 5, $\mathcal{C}'$ is stable. It follows that $\mathcal{C}$ is also stable. $\square$

With a minor extension of the same method, we can also prove the more general result for classes of structures of an arbitrary fixed signature.

5 Independence property

A first-order formula $\varphi(\bar{x}, \bar{y})$ is said to have the *independence property* with respect to $\mathcal{C}$ if it has the *n-independence property* for all n, i.e. if for every n there exist a structure $M \in \mathcal{C}$ and tuples $\bar{a}_0, \ldots, \bar{a}_{n-1} \in M$ and $\bar{b}_J \in M$ for all $J \subseteq \{0, 1, \ldots, n-1\}$ such that $M \models \varphi(\bar{a}_i, \bar{b}_J)$ holds if and only if $i \in J$. $\mathcal{C}$ is said to be *dependent* or to have *NIP* if no formula has the independence property with respect to $\mathcal{C}$. One can show that $\varphi(\bar{u}, \bar{v})$ has the independence property if and only if the 'opposite formula' $\varphi(\bar{v}, \bar{u})$ (i.e. really the same formula, but listing the variables differently) has it. It is easy to see that every formula with the (n-)independence property has the (n-)order property. Therefore every stable class is dependent. See [1] for more on the independence property and its relation to the order property. A related notion is the VC dimension from statistical learning theory. Applied to a formula, it is essentially the greatest n such that the formula has the n-independence property.[10]

Like stability, the independence property comes from stability theory [15, 6], and is originally only defined for first-order theories. Again, a single structure M is called dependent if $\{M\}$ is dependent or, equivalently, if the class of all structures elementarily equivalent to M is dependent.

Lemma 8. *Let $\mathcal{C}$ be a subgraph-closed class of graphs. If $\mathcal{C}$ is dependent, then $\mathcal{C}$ is superflat.*

Proof. Suppose $\mathcal{C}$ is not superflat, i.e. for some r, every K_m^r occurs as a subgraph of a member of $\mathcal{C}$. Since the following graph A_m is a subgraph of $K_{m+2^m}^r$, it also occurs as a subgraph of a member of $\mathcal{C}$, hence is itself a member of $\mathcal{C}$ (up to isomorphism). A_m has vertices $a_0, a_1, \ldots, a_{m-1}$ and b_J for each $J \subseteq \{0, 1, \ldots, m-1\}$ as well as additional vertices that appear in the following. For any $i \in \{0, 1, \ldots, m-1\}$ and any $J \subseteq \{0, 1, \ldots, m-1\}$ such that $i \in J$, there is a path of length $r+1$ from a_i to b_J. The interior parts of these paths are pairwise disjoint and disjoint from the set of vertices a_i and b_J. There are no further vertices or edges.

Let $\varphi(x, y)$ be the formula that says that there is a path of length $r+1$ from x to y. Since $A_m \models \varphi(a_i, b_J)$ if and only if $i \in J$, the family of graphs A_m witnesses that $\varphi(x, y)$ has the independence property with respect to $\mathcal{C}$. So $\mathcal{C}$ is not dependent. □

It is an open question whether Lemma 8 can be generalised to relational structures of finite signatures.

We will call a class $\mathcal{C}$ of relational structures *monotone* if whenever $M \to N$ is an injective homomorphism and $N \in \mathcal{C}$, we also have $M \in \mathcal{C}$. In other words, a monotone class is closed under isomorphism and 'non-induced substructures', the natural generalisation of non-induced subgraphs to arbitrary signatures. Putting all the previous results together, we have the following theorem.

Theorem 9. *Let $\mathcal{C}$ be a monotone class of coloured digraphs of a fixed finite signature. The following conditions are equivalent.*

1. *$\underline{\mathcal{C}}$ is nowhere dense.*
2. *$\underline{\mathcal{C}}$ is superflat.*
3. *$\mathcal{C}$ is stable.*
4. *$\underline{\mathcal{C}}$ is stable.*
5. *$\mathcal{C}$ is dependent.*
6. *$\underline{\mathcal{C}}$ is dependent.*

Proof. The first two conditions are equivalent by Remark 2 and imply the third by Theorem 7. $3 \Rightarrow 5$ and $4 \Rightarrow 6$ because every stable formula is dependent. $3 \Rightarrow 4$ and $5 \Rightarrow 6$ because there is an interpretation $\underline{\mathcal{C}} \to \mathcal{C}$ by finiteness of the signature. Finally, $6 \Rightarrow 1$ by Lemma 8 because $\underline{\mathcal{C}}$ is closed under subgraphs by monotonicity of $\mathcal{C}$. □

As a corollary to the proof, we see that to check stability or dependence of a monotone class of coloured digraphs it is sufficient to look at formulas of the form $\varphi(x, y)$ with single variables x and y. This is not true in general.

The condition that $\mathcal{C}$ be monotone is crucial. The class of all complete graphs is stable but not nowhere dense / superflat. It is also not hard to code the class of linear orders in a class of graphs that is dependent but not stable. Further it is crucial for the equivalence of stability and NIP that $\mathcal{C}$ is a class of relational structures (i.e. the signature does not contain function symbols), since the class of all linear orders coded by the binary function min is dependent but not stable. Also note that the underlying graph of a structure with a binary function symbol is always complete. Finally, in a signature with infinitely many binary relation symbols, $\underline{\mathcal{C}}$ may not be interpretable in $\mathcal{C}$, and in fact a monotone class $\mathcal{C}$ of such a signature may be stable even though $\underline{\mathcal{C}}$ is not.

We obtain a corollary in computational learning theory, which can be seen as the analogue to an earlier result by Grohe and Turán, connecting finite VC dimension of the monadic second order formulas on a subgraph-closed class with bounded tree-width [7].

Corollary 10. *Let $\mathcal{C}$ be a subgraph-closed class of graphs. Every concept class definable in first-order logic on $\mathcal{C}$ has bounded sample complexity in the PAC model, if and only if $\mathcal{C}$ is nowhere dense.*

This immediately follows from Theorem 9 using the standard connection between VC dimension and sample complexity in the PAC model of computational learning theory (see e. g. [7]).

6 Conclusion

In this paper, we proved that for a class of graphs $\mathcal{C}$ that is closed under subgraphs, $\mathcal{C}$ being nowhere dense is equivalent to $\mathcal{C}$ being stable, and to $\mathcal{C}$ being NIP (i.e. first-order definable concept classes on C being PAC learnable), and we extended this result to classes of coloured digraphs. We leave the following open problem.

Open problem. Can Theorem 9 be generalised to arbitrary relational structures with finite signatures?

Altogether, we have seen that tameness notions from combinatorial graph theory, finite model theory and stability theory can be compared for classes of graphs, so long as they are closed under subgraphs. The latter restriction is a rather natural one in the first two fields, but severe and unnatural from the point of view of stability theory, of which the observed collapse of stability and NIP may be a symptom. This will probably make a transfer of ideas from stability theory to the other fields more straightforward than in the other direction. It remains to be seen to what extent parts of stability theory (indiscernibles, forking or splitting, etc.) can be generalised to the new context and whether they are of any relevance to the algorithmically oriented fields.

Using work of Herre, Mekler and Smith [8] and a straightforward generalisation of the model theoretic notion of strong stability into our context, one can show that a class $\mathcal{C}$ of coloured digraphs such that $\mathcal{C}$ omits a finite *topological* minor, is *strongly* stable. Similar statements concerning further model theoretic dividing lines such as superstability or simplicity would be interesting, but for these notions the generalisation to arbitrary classes of structures seems to be less straightforward.

Finally, we hope that bringing together the tools from the different fields will make it easier to find a unifying combinatorial explanation for the algorithmically tame (or wild) behaviour of many graph classes.

References

[1] Hans Adler. Introduction to theories without the independence property. *Arch. Math. Log.* accepted.

[2] Anuj Dawar. Finite model theory on tame classes of structures. In Ludek Kučera and Antonín Kučera, editors, *MFCS*, volume 4708 of *Lect. Notes Comput. Sc.*, pages 2–12. Springer, 2007.

[3] Anuj Dawar. Homomorphism preservation on quasi-wide classes. *J. Comput. Syst. Sci.*, 76(5):324–332, 2010.

[4] Anuj Dawar and Stephan Kreutzer. Domination problems in nowhere-dense classes. In Ravi Kannan and K. Narayan Kumar, editors, *FSTTCS*, volume 4 of *LIPIcs*, pages 157–168. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2009.

[5] Reinhard Diestel. *Graph Theory*, volume 173 of *Grad. Texts in Math.* Springer, 2005.

[6] Doug Ensley and Rami Grossberg. Finite models, stability, and Ramsey's theorem, 1996.

[7] Martin Grohe and György Turán. Learnability and definability in trees and similar structures. *Theory Comput. Syst.*, 37(1):193–220, 2004.

[8] Heinrich Herre, Alan H. Mekler, and Kenneth W. Smith. Superstable graphs. *Fund. Math.*, 118:75–79, 1983.

[9] Wilfrid Hodges. *A Shorter Model Theory.* Cambridge University Press, 1997.

[10] M. Chris Laskowski. Vapnik-Chervonenkis classes of definable sets. *J. London Math. Soc. (2)*, 45:377–384, 1992.

[11] Jaroslav Nesetril and Patrice Ossona de Mendez. *Sparsity - Graphs, Structures, and Algorithms*, volume 28 of *Algorithms and combinatorics*. Springer, 2012.

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

[13] Jaroslav Nešetřil and Patrice Ossona de Mendez. First order properties on nowhere dense structures. *J. Symb. Logic*, 75:868–887, 2010.

[14] Jaroslav Nešetřil and Patrice Ossona de Mendez. On nowhere dense graphs. *Eur. J. Comb.*, 32(4):600–617, 2011.

[15] Saharon Shelah. *Classification Theory and the Number of Non-Isomorphic Models*. Stud. Logic Found. Math. North-Holland, 2nd edition, 1990.

[16] Vladimir N. Vapnik and Alexey Y. Chervonenkis. On the uniform convergence of relative sequences of events to their probabilities. *Theory Probab. Appl.*, 16:264–280, 1971.

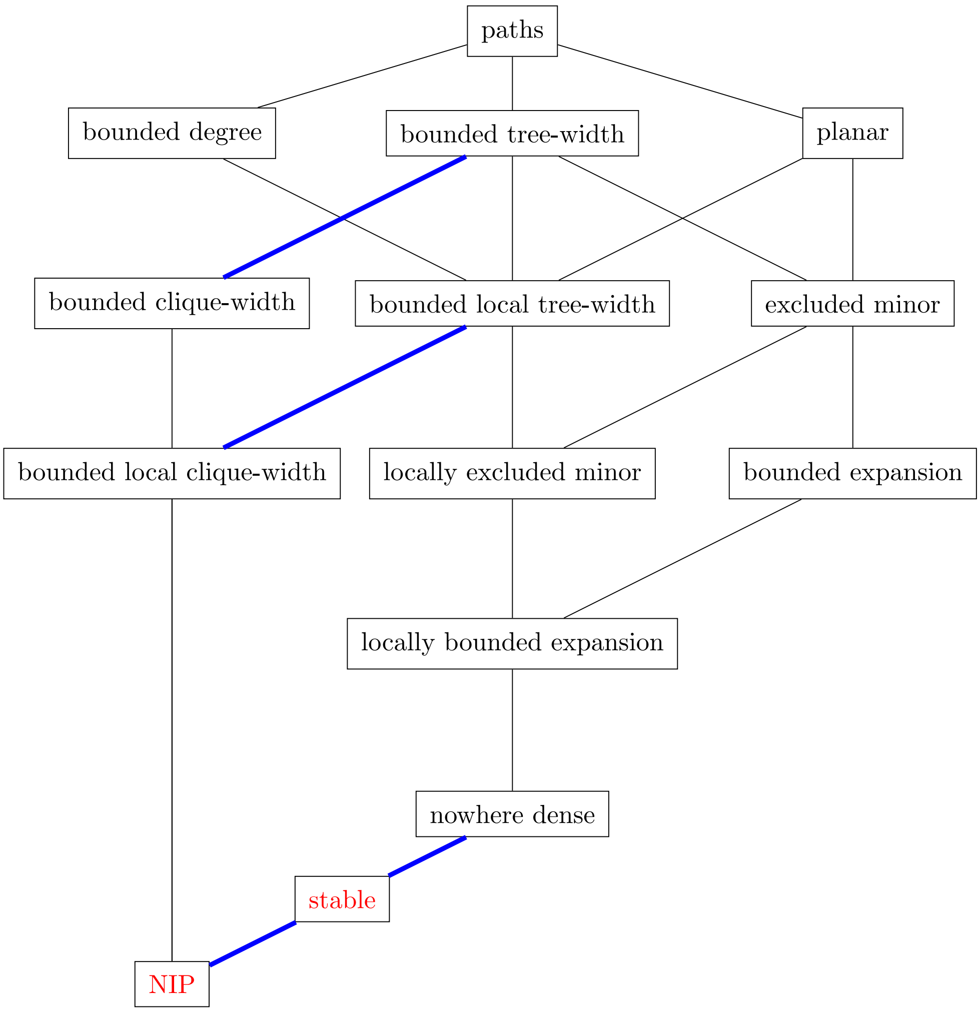

Figure 2: Inclusion diagramm of graph classes. The lines indicate inclusions (top-down). A bold (blue) line connects graph classes that coincide if they are subgraph-closed. The three lower classes are studied in this paper. For the other graph classes, see e. g. [7, 14]. Note that the class of graphs of clique-width 3 is not stable.