A New Perspective on FO Model Checking of Dense Graph Classes

J. Gajarský and P. Hliněný and J. Obdržálek
Masaryk University, Brno
{gajarsky,hlineny,obdrzalek}@fi.muni.cz

D. Lokshtanov
University of Bergen
daniello@ii.uib.no

M. S. Ramanujan
TU Wien
ramanujan@ac.tuwien.ac.at

Abstract

We study the FO model checking problem of dense graph classes, namely those which are FO-interpretable in some sparse graph classes. Note that if an input dense graph is given together with the corresponding FO interpretation in a sparse graph, one can easily solve the model checking problem using the existing algorithms for sparse graph classes. However, if the assumed interpretation is not given, then the situation is markedly harder.

In this paper we give a structural characterization of graph classes which are FO interpretable in graph classes of bounded degree. This characterization allows us to efficiently compute such an interpretation for an input graph. As a consequence, we obtain an FPT algorithm for FO model checking of graph classes FO interpretable in graph classes of bounded degree. The approach we use to obtain these results may also be of independent interest.

Categories and Subject Descriptors F.4.1 [*Mathematical Logic and Formal Languages*]: Mathematical Logic—Model theory

General Terms FO Logic, Model-Checking, Logic Interpretations, Sparse Graph Classes, Parameterized Complexity

Keywords FO Logic, Model-Checking, Logic Interpretations, Sparse Graph Classes, Parameterized Complexity

1. Introduction

Algorithmic metatheorems are theorems stating that all problems expressible in a certain logic are efficiently solvable on certain classes of (relational) structures, e.g. on finite graphs. Note that the model checking problem for *first-order logic* – given a graph G and an FO formula ϕ we want to decide whether G satisfies ϕ (written as $G \models \phi$) – is trivially solvable in time $|V(G)|^{\mathcal{O}(|\phi|)}$. "Efficient solvability" hence in this context often means *fixed-parameter tractability* (FPT); that is, solvability in time $f(|\phi|) \cdot |V(G)|^{\mathcal{O}(1)}$ for some computable function f.

In the past two decades algorithmic metatheorems for FO logic on sparse graph classes received considerable attention. After the result of Seese [16] establishing fixed-parameter tractability of FO model checking on graphs of bounded degree there followed a series of results [3, 5, 6] establishing the same result for increasingly rich sparse graph classes. This line of research culminated in the result of Grohe, Kreutzer and Siebertz [11], who proved that FO model checking is FPT on nowhere dense graph classes.

The result of Grohe, Kreutzer and Siebertz [11] is essentially the best possible of its kind, in the following sense: If a graph class $\mathcal{D}$ is monotone (i.e., closed on taking subgraphs) and not nowhere dense, then the FO model checking problem on $\mathcal{D}$ is as hard as that on all graphs. Possible ways to continue the research into algorithmic metatheorems for FO logic include the following two.

First, one can study relational structures other than graphs. This line of research has recently been initiated by Bova, Ganian and Szeider [1], who gave an FPT algorithm for existential FO model checking on partially ordered sets of bounded size of a maximum antichain. Shortly after followed the result of Gajarský et al. [8], who extended [1] to full FO. Apart from these results, very little is known and it remains to be seen what other types of structures and their parameterizations admit fast FO model checking algorithms. Second, one may consider metatheorems for FO logic on classes of graphs which are not sparse. Again, little is known along this line of research. One can mention the result of Ganian et al. [10] establishing that certain subclasses of interval graphs admit an FPT algorithm for FO model checking. Besides, the aforementioned result of [8] can also be seen as a result about dense (albeit directed) graphs, and [8] actually happens to imply the result of [10].

We would like to initiate a systematic study of dense graph classes for which the FO model checking problem is efficiently solvable. It appears that a natural way to arrive at new graph classes admitting FPT algorithms for FO model checking, is by *means of interpretation*. In a simplified setting (although, see Section 4.3 for the more general case) – given a graph G and an FO formula $\psi(x, y)$ with two free variables, we can define a graph $H = I_\psi(G)$ on the same vertex set as G and the edge set determined by $\psi(x, y)$: a pair of distinct vertices u, v is an edge of H iff $G \models \psi(u, v)$ or $G \models \psi(v, u)$. We then say that H is interpreted in G using ψ. A graph class $\mathcal{D}$ is interpretable in a graph class $\mathcal{C}$ if there exists an FO formula $\psi(x, y)$ such that every member of $\mathcal{D}$ is interpreted in some member of $\mathcal{C}$ using ψ.

In this context we ask the following question:

Question 1.1. *Let $\mathcal{C}$ be a graph class admitting an FPT algorithm for FO model checking, and $\mathcal{D}$ be a graph class interpretable in $\mathcal{C}$. Does there exist an FPT algorithm for FO model checking on $\mathcal{D}$?*

It might seem that a definite easy answer is 'yes', based on the following natural property of interpretations: if $H \in \mathcal{D}$ is interpreted in $G \in \mathcal{C}$ using formula $\psi(x, y)$, and our question is to decide whether $H \models \phi$, it is a standard routine to construct a sentence ϕ' such that $H \models \phi$ if and only if $G \models \phi'$. Then $G \models \phi'$ is decided by the FPT algorithm given for $\mathcal{C}$. However, the difficulty lies in the fact that our inputs come from $\mathcal{D}$, without any reference to the respective members of $\mathcal{C}$ in which they are interpreted. Even if the interpretation formula $\psi(x, y)$ is fixed and known beforehand, we have generally no efficient way of obtaining the respective member $G \in \mathcal{C}$ for an input $H \in \mathcal{D}$. Thus, Question 1.1 can be reduced to the following:

Question 1.2. *Let $\mathcal{C}, \mathcal{D}$ be graph classes such that $\mathcal{D}$ is interpretable in $\mathcal{C}$. Does there exist an integer s and a polynomial-time algorithm $\mathcal{A}$ such that; given $H \in \mathcal{D}$ as input, $\mathcal{A}$ outputs $G \in \mathcal{C}$ and an FO formula $\psi(x, y)$ of size at most s such that H is interpreted in G using ψ ?*

An answer to Question 1.2 is far from being obvious. Take, for example, the following particular FO interpretation: A graph H is the *square* of a graph G if the edges of H are those pairs of vertices which are at distance at most 2 in G. Then the problem; given H find G such that H is the square of G, is NP-hard [14]. Another such negative example, specifically tailored to our settings, is discussed in Section 4.4. These examples show that it is important to choose a suitable interpretation formula ψ (avoiding the hard cases) in an answer to Question 1.2.

Our contribution. We answer both Questions 1.1 and 1.2 in the positive for the case when $\mathcal{C}$ is a class of graphs of bounded degree.

We first define near-uniform graph classes, based on a new notion of near-k-twin relation, which generalizes the folklore twin-vertex relation, and is related also to the neighbourhood diversity parameter of [12]. The idea behind this approach is to classify pairs of vertices which have almost the same adjacency to the rest of the graph. The approach seems promising and may be of independent use in further investigation of well structured dense graph classes.

We then give an efficient FO model checking algorithm for the near-uniform graph classes. This algorithm is based upon the above idea of interpretation; briefly, given a graph H we use the near-k-twin relation for a suitable value of k to partition the vertex set of H and to find a bounded degree graph G, such that H is interpreted in G using a universal formula ψ depending only on the class in question. Then we employ the aforementioned algorithm of Seese [16].

In the second half of the paper we show that the concept of near-uniform graph classes is robust and sufficiently rich in content, by proving that the near-uniform graph classes are exactly the classes which are FO interpretable in graphs of bounded degree. At this place we remark that properties of graphs which are FO interpretable in graphs of bounded degree have already been studied, e.g., by Dong, Libkin and Wong in [4] in a different context, but those previous results do not imply our conclusions. We finish by sketching some interesting open directions for future research.

2. Preliminaries and outline

In this section we define and discuss interpretations and afterwards give a brief exposition of ideas leading to our results. We start by explaining the core ideas behind our approach to analysing dense graphs and then we sketch the how interpretations are combined with our approach to dense graphs to obtain the results presented in Sections 3 and 4.

Graph theory. We work with finite simple undirected graphs and use standard graph theoretic notation.

Interpretations. We recall the definitions of FO interpretability for graphs and graph classes. In order to simplify our exposition and proofs we use a slightly simplified version of interpretations. However, as we argue later (Section 4.3), this does not really lead to a loss of generality in our case.

Let $\psi(x, y)$ be an FO formula with two free variables over the language of (possibly labelled) graphs such that for any graph and any u, v it holds that $G \models \psi(u, v) \Leftrightarrow G \models \psi(v, u)$ and $G \not\models \psi(u, u)$, i.e. the relation on $V(G)$ defined by the formula is symmetric and irreflexive. From now on we will assume that formulas with two free variables are symmetric and irreflexive. Given a graph G, the formula $\psi(x, y)$ maps G to a graph $H = I_\psi(G)$ defined by $V(H) = V(G)$ and $E(H) = \{\{u, v\} \mid G \models \psi(u, v)\}$. We then say that the graph H is *interpreted* in G. Notice that even though the graph G can be labelled, our graph H is not. This is to simplify our arguments – nevertheless, one may easily inherit labels from G to H if needed.

In the rest of the paper, whenever we consider graphs G and H in context of interpretations, graph G will be the graph in which we are interpreting, and graph H will be the "result" of the interpretation.

The notion of interpretation can be extended to graph classes – to a graph class $\mathcal{C}$ the formula $\psi(x, y)$ assigns the graph class $\mathcal{D} = I_\psi(\mathcal{C}) = \{H \mid H = I_\psi(G), G \in \mathcal{C}\}$. We say that a graph class $\mathcal{D}$ is *interpretable* in a graph class $\mathcal{C}$ if there exists formula $\psi(x, y)$ such that $\mathcal{D} \subseteq I_\psi(\mathcal{C})$. Note that we do not require $\mathcal{D} = I_\psi(\mathcal{C})$, we just want every graph from $\mathcal{D}$ to have a preimage in $\mathcal{C}$.

Interpretations are useful for defining new graphs from old using logic (again, we think of H as a result of application of ψ to G), but can also be used to evaluate formulas on H quickly, provided that we have a fast algorithm to evaluate formulas on G. Let $H = I_\psi(G)$, let θ be a sentence and let θ' be a sentence obtained from θ by replacing every occurrence of the atom edge(x, y) by $\psi(x, y)$. Then $H \models \theta \iff G \models \theta'$.

Note that in our definition of interpretation we require that $V(H) = V(G)$ and use the formula $\psi(x, y)$ to define the edges of H, while the usual definition employs a pair of formulas $\nu(x)$ and $\mu(x, y)$ so that $V(H) = \{u \in V(G) \mid G \models \nu(u)\}$ and $E(H) = \{\{u, v\} \subseteq V(H) \mid G \models \mu(u, v)\}$. Nevertheless, this simplification does not lead to a loss of generality: First it is easy to see that our version of interpretation is just an instance of the ordinary one, by taking $\mu(x, y) = \psi(x, y)$ and setting $\nu(x) =$ true. As for the opposite direction, we show in Section 4.3 that graph classes obtained from graphs of bounded degree using ordinary interpretations (using $\nu(x)$ and $\mu(x, y)$) can be equivalently characterized using only $\psi(x, y)$.

2.1 Locality, indistinguishability, and the new approach

The existing FPT algorithms for FO model checking of sparse graph classes we mentioned at the beginning of Section 1 rely heavily on the use of locality of FO logic – the problem of evaluating FO formulas can be reduced to evaluating *local* FO formulas (cf. Gaifman theorem [7], also in Section 4). This, together with the fact that in sparse graphs it is possible to evaluate local formulas efficiently, made the locality-based approach suitable for studying FO logic on sparse graphs. The problem with using this approach for dense graphs is obvious – in a dense graph the whole graph can be in the 1-neighbourhood of a single vertex[1]. This makes evaluating local formulas around such a vertex expensive (from the FPT perspective), because this amounts to evaluating them on the whole graph.

An alternative approach to FO model checking is based on the concept of vertex indistinguishability. This approach can be used for dense graphs, but is a bit too limited in its scope. The key notion here is that of twin vertices – two vertices of a graph G are *twins* if they have the same neighbourhood. The fact that two vertices u, v are twins means that they behave the same with respect to any other vertex in a graph. Consequently, no FO formula can

[1] This is also true for some sparse graphs, say stars, but we hope that it is clear that for dense graphs this can cause substantial problems.

distinguish between u and v. It is not hard to see that the twin relation is an equivalence on the vertex set of a graph. The number of equivalence classes of this relation is called the *neighbourhood diversity* [12] of a graph and graph classes of bounded neighbourhood diversity admit a very simple FPT algorithm for FO model checking. However, as already mentioned, the problem with this approach is that it is too restrictive – even such simple graph classes as paths have unbounded neighbourhood diversity.

Our approach is based on observing that the locality-based approach, when used on sparse graphs, is in its essence based on indistinguishability of vertices. For example, in a graph of bounded degree, all vertices behave in the same way with respect to the rest of the vertex set (they are non-adjacent to it), with only a few exceptions (their 1-neighbourhoods). In other words, any two vertices have *almost* the same neighbourhood. This leads to the relaxation of the notion of twin vertices. We say that two vertices are *near-k-twins* if their neighbourhoods differ in at most k vertices. To see how this notion works around the issues with locality and indistinguishability explained above, let us consider the near-k-twin relation on the class $\mathcal{G}_d$ of graphs of degree at most d and on the class $\overline{\mathcal{G}_d}$ of its complements. On every graph from these graph classes, the near-$2d$-twin relation is an equivalence with one class. Graphs from $\overline{\mathcal{G}_d}$ are dense and some of them contain universal vertices. Moreover, the class of all paths is a subset of $\mathcal{G}_2$.

It is important to note that, unlike the ordinary twin relation, the near-k-twin relation is not automatically guaranteed to be an equivalence (this depends heavily on specific G and k). However, when it indeed is an equivalence, it enjoys some desirable structural properties. This leads us to studying graph classes such that for each graph from these classes there exists a small k such that the near-k-twin relation is an equivalence with a small number of classes – the *near-uniform* graph classes. The precise definition of near-uniformity, as well as the details of the above mentioned structural properties are explained in the first part of Section 3, where we also show how to exploit these properties to obtain an FPT algorithm for FO model checking on near-uniform graph classes.

2.2 Interpretability in graphs of bounded degree

The graph classes which are interpretable in graphs of bounded degree are studied in Section 4. Their characterization relies on a simple corollary of Gaifman's locality theorem: For a graph G and two vertices $u, v \in V(G)$ which are far apart form each other, the validity formula $\psi(u, v)$ depends only on formulas with one free variable (up to the quantifier rank q, which depends on ψ) valid on u and v (i.e. its logical q-types). This in turn means that when the formula $\psi(u, v)$ is used for interpretation (to obtain the graph H from a graph G of degree at most d) and vertices u and u' satisfy the same formulas with one free variable (again, up to the quantifier rank q), u and u' will be adjacent to the same vertices in the resulting graph, except for a small number of vertices which were in their respective r-neighbourhoods in graph G (here r also depends on $\psi(x, y)$). Any two vertices of the same q-type will therefore be near-k-twins for $k = 2 \cdot d^r$. Since the relation "being of the same q-type" is an equivalence with a bounded number of classes, this could lead one to believe that the near-k-twin relation (for a suitably chosen k) is an equivalence with a bounded number of classes (independent of G) for any graph from a graph class interpretable in a class of graphs of bounded degree. This, however, is not true – it can happen that some vertices u and v of different q-types can be near-k-twins and vertex w of yet different q-type can be near-k-twin of v and not a near-k twin of u, thus failing the transitivity.

The desired characterization of graph classes interpretable in classes of graphs of bounded degree therefore has a bit more complicated form: A class $\mathcal{D}$ is interpretable in a class of graphs of bounded degree if there exist k_0 and p such that for every $H \in \mathcal{D}$ there exists $k \in \{0, \ldots, k_0\}$ such that near-k-twin relation is an equivalence on $V(H)$ with at most p classes – this characterization thus precisely coincides with near-uniformity. Notice that this, in turn, also implies the existence of an FPT model checking algorithm for graph classes interpretable in classes of graphs of bounded degree.

3. Near-k-twins and FO model checking

In this section we extend the concept of indistinguishability of twin vertices and show how this can lead to an efficient FO model checking algorithm. This constitutes the main algorithmic contribution of the paper. We start by giving a definition of the near-k-twin relation and outlining the model checking algorithm based on this, while we postpone further details and a proof of Theorem 3.3 to the rest of this section.

We begin by clarifying the terminology. We assume that 0 is a natural number, i.e. $0 \in \mathbb{N}$. Let $X \triangle Y$ denote the symmetric difference of two sets. For a graph G and a vertex $v \in V(G)$, we define the *neighbourhood* of v as $N^G(v) = \{w \in V(G) \mid \{v, w\} \in E(G)\}$. If the graph G is clear from the context, we write just $N(v)$. Note that, by definition, $v \notin N(v)$.

A useful concept in graph theory is that of twin vertices. Two vertices $u, v \in V(G)$ are called *false twins* if $N(u) = N(v)$, and they are *true twins* if $N(u) \cup \{u\} = N(v) \cup \{v\}$. We actually follow the concept of false twins, which better suits our purposes, in the next definition.

Definition 3.1 (near-k-twin)**.** *For a graph G and $k \in \mathbb{N}$, the* near-k-twin relation of G *is the relation ρ_k on $V(G)$ defined by* $(u, v) \in \rho_k \iff |N(u) \triangle N(v)| \le k$.

Considering, e.g., k a small parameter and G a large graph then, intuitively, two vertices of G are near-k-twins if they have "almost the same" neighbourhood. This relation, unlike the ordinary twin relations on graph vertices, does not always "behave nicely"; in particular, ρ_k may not be an equivalence relation (see e.g. the examples below). On the other hand, if the near-k-twin relation is an equivalence of bounded index, then we can use it to decompose the vertex set of the graph G and efficiently find an interpretation of G in a graph of bounded degree. This leads to the following.

Definition 3.2 (near-uniform)**.** *A graph G is* (k_0, p)-near-uniform *if there exists $k \le k_0$ for which near-k-twin relation of H is an equivalence of index at most p.*
A graph class $\mathcal{D}$ is (k_0, p)-near-uniform *if every member of $\mathcal{D}$ is (k_0, p)-near-uniform, and $\mathcal{D}$ is* near-uniform *if there exist integers k_0, p such that $\mathcal{D}$ is (k_0, p)-near-uniform.*

Our algorithm for near-uniform graph classes can be summarized in the following three steps.

The FO model checking algorithm. Given is graph H from a (k_0, p)-near-uniform graph class $\mathcal{D}$ and an FO formula ϕ. Perform the following steps:

1. For each $k := 0, 1, \ldots, k_0$; compute the near-k-twin relation ρ_k of H, and check whether ρ_k is an equivalence of index at most p. This test has to succeed for some value of k.
2. Compute a universal formula $\psi(x, y)$ depending on k_0 and p, and the graph G_H depending on H and k found in step 1, such that $H = I_\psi(G_H)$ and the vertex degrees in G_H are at most $2k_0p$ (Theorem 3.7).
3. Run the algorithm of [16] for FO model checking on graphs of bounded degree on G_H and the formula ϕ', where ϕ' is

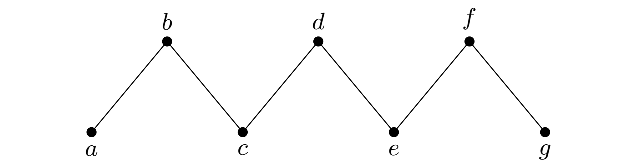

Figure 1. An example. The near-2-twin relation ρ_2 of this path includes pairs (b, d) and (d, f) but not (b, f), and so ρ_2 is not an equivalence. On the other hand, ρ_1 is an equivalence on this path and its near-1-twin classes are $\{a, c\}, \{e, g\}, \{b\}, \{d\}, \{f\}$.

obtained from ϕ by replacing every occurrence of $edge(z, z')$ with $\psi(z, z')$.

Theorem 3.3. *Let $\mathcal{D}$ be a (k_0, p)-near-uniform graph class for some $k_0, p \in \mathbb{N}$. Then the FO model checking problem in $\mathcal{D}$ is fixed-parameter tractable when parameterized by the formula size, i.e., solvable in time $f(|\phi|) \cdot |V(G)|^{\mathcal{O}(1)}$ for a computable function f and input G, ϕ.*

3.1 Properties of the near-k-twin relation

To give details of the algorithm and to prove Theorem 3.3, we study some structural properties of graphs for which the near-k-twin relation is actually an equivalence. To simplify the discussion, we use the following notation. If ρ_k of Definition 3.1 is an equivalence relation, then we call ρ_k the *near-k-twin equivalence* of G, and the equivalence classes of ρ_k the *near-k-twin classes of G*.

For example, take a class $\mathcal{G}_d$ of the graphs of maximum degree at most d, and let $k = 2d$. Then the near-k-twin relation ρ_k is a trivial equivalence of index one (i.e., with one class) for every graph from $\mathcal{G}_d$. The same holds for the class $\overline{\mathcal{G}}_d$ of the complements of graphs of $\mathcal{G}_d$. Another sort of examples comes, say, with a class $\overline{\mathcal{B}}_d$ of the graphs obtained from complete bipartite graphs by subtracting a subgraph of degrees at most d. For $k = 2d$ and every graph of $\overline{\mathcal{B}}_d$, the near-k-twin relation ρ_k is an equivalence of index at most two.

On the other hand, we can easily see that the near-2-twin relation of, e.g., a path of length 6 is not an equivalence; Figure 1. Even more, examples such as that of Figure 1 show that, having a near-k-twin equivalence for some k, does not imply that the near-k'-twin relation is an equivalence for $k' > k$. That is why we cannot simply use one universal value of k in Definition 3.2.

As outlined above in the algorithm, our key step is to show that all near-uniform graph classes are FO interpretable in graph classes of bounded degree. For this we show that for any two large enough equivalence classes of a near-k-twin equivalence, it holds that every vertex from one class is connected to almost all or to almost none vertices of the other class and vice versa. More precisely:

Lemma 3.4. *Let $k \geq 1$ and G be a graph such that the near-k-twin relation ρ_k of G is an equivalence on $V(G)$. Let U and V be two near-k-twin classes of G with at least $4k + 2$ vertices each (it may be $U = V$). Then for every $v \in V$ we have*

$$\min\{|U \cap N(v)|,\ |U \setminus N(v)|\} \leq 2k.$$

Note that the claim of Lemma 3.4 universally holds only when both U and V are sufficiently large. A counterexample with small U is a graph consisting of $U = \{u\}$ and V inducing a large clique, such that u is connected to half of the vertices of V. For this graph the near-1-twin classes are exactly U and V, but both $|V \cap N(u)|$ and $|V \setminus N(u)|$ are unbounded.

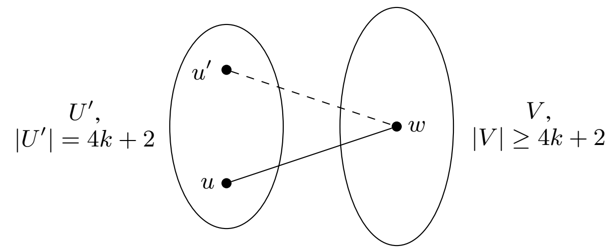

Figure 2. An illustration; counting the pairs $(w, \{u, u'\})$ such that $w \in V$, $u, u' \in U'$ in the proof of Lemma 3.4, in case $U \neq V$.

Proof. For $x \in V(G)$ and $A \subseteq V(G)$, let $\alpha^A(x) = \min\{|N(x) \cap A|,\ |A \setminus N(x)|\}$. Thus to prove the lemma we need to show that $\alpha^U(v) \leq 2k$ for $v \in V$.

Towards a contradiction assume $\alpha^U(v) \geq 2k + 1$ for some $v \in V$. Clearly, there is a subset $U' \subseteq U$ such that $|U'| = 4k + 2$ and $\alpha^{U'}(v) \geq 2k + 1$, too. Since $|N(w) \triangle N(v)| \leq k$ for any $w \in V$ by the definition of ρ_k, we also get $\alpha^{U'}(w) \geq \alpha^{U'}(v) - k \geq 2k + 1 - k = k + 1$ for all $w \in V$.

We are going to count the number D of pairs $(w, \{u, u'\})$ such that $w \in V$, $u, u' \in U'$ are distinct vertices and exactly one of wu, wu' is an edge of G. See Figure 2. On the one hand, for any fixed $u, u' \in U'$, every w forming such a desired pair $(w, \{u, u'\})$ belongs to $N(u) \triangle N(u')$ and so we have got an upper bound

$$\begin{aligned} D &\leq \sum_{\{u,u'\} \in \binom{U'}{2}} |N(u) \triangle N(u')| \leq \\ &\leq \binom{|U'|}{2} \cdot k = \binom{4k+2}{2} \cdot k < 3k^2(4k+2)\,, \end{aligned} \tag{1}$$

where $|N(u) \triangle N(u')| \leq k$ holds by the definition of ρ_k.

On the other hand, we may fix $w \in V$ and count the number of unordered pairs $u, u' \in U' \setminus \{w\}$ such that exactly one of wu, wu' is an edge of G; this number is equal to $|N(w) \cap U'| \cdot |U' \setminus N(w)| = \alpha^{U'}(w) \cdot \big(|U'| - \alpha^{U'}(w)\big)$ if $w \notin U'$, and to $\alpha^{U'}(w) \cdot \big(|U'| - 1 - \alpha^{U'}(w)\big)$ or $(\alpha^{U'}(w) - 1) \cdot \big(|U'| - \alpha^{U'}(w)\big)$ if $w \in U'$. Therefore,

$$\begin{aligned} D &\geq \sum_{w \in V} \big(\alpha^{U'}(w) - 1\big) \cdot \big(|U'| - 1 - \alpha^{U'}(w)\big) \\ &\geq \sum_{w \in V} (k + 1 - 1)(4k + 2 - 1 - k - 1) \\ &= |V| \cdot 3k^2 \geq 3k^2(4k + 2) \end{aligned} \tag{2}$$

since we have got $\alpha^{U'}(w) \geq k + 1$ and $|V| \geq 4k + 2 = |U'|$.

Now, (1) and (2) are in a contradiction, and hence the sought conclusion follows. $\square$

Corollary 3.5. *Let U and V be the two classes of Lemma 3.4 such that $|U|, |V| \geq 5k + 1$. Then exactly one of the following two possibilities holds:*

a) every vertex of U is connected to at most $2k$ vertices of V and every vertex of V is connected to at most $2k$ vertices of U, or

b) every vertex of U is connected to all but *at most $2k$ vertices of V and every vertex of V is connected to* all but *at most $2k$ vertices of U.*

Proof. We first show that either

- every vertex of U is connected to at most $2k$ vertices of V, or

- every vertex of U is connected to all but $2k$ vertices of V.

Indeed, for any vertex $v \in U$ taken separately, only one of these cases can happen since $|V| > 4k$, and one of these cases has to happen by Lemma 3.4. Assume that there exist $v, w \in U$ with v having at most $2k$ neighbours in V while w is connected to all but at most $2k$ vertices of V. Then $|N(v) \triangle N(w)| \geq |V| - 2k - 2k \geq k+1$, contradicting the definition of ρ_k.

To finish the proof, we have to show the the following case (relevant if $U \neq V$) is impossible: every vertex of U connected to at most $2k$ vertices of V and every vertex of V connected to *all but* at most $2k$ vertices of U. In the argument we count the total number of edges between U and V; it would be at most $2k \cdot |U|$ and, at the same time, at least $(|U| - 2k) \cdot |V|$. Though, the difference between these lower and upper estimates is

$$\begin{aligned} & (|U| - 2k) \cdot |V| \;-\; 2k \cdot |U| \\ = \;& |U| \cdot |V| - 2k \cdot (|U| + |V|) \\ = \;& \big(|U| - 4k\big)\big(|V| - 4k\big) + 2k\big(|U| + |V|\big) - 16k^2 \\ > \;& k \cdot k + 2k(5k + 5k) - 16k^2 = 5k^2 > 0, \end{aligned} \tag{3}$$

a contradiction, thus finishing the whole proof. ☐

Remark 3.6. Note that Corollary 3.5 still applies if $U = V$. I.e., for a single near-k-twin equivalence class U with $|U| > 5k + 1$ either

a) every vertex of U has at most $2k$ neighbours in U, or
b) every vertex of U has at least $|U| - 2k$ neighbours in U.

3.2 From near-k-twins to bounded degree

Here we present the core of our algorithm – a procedure which, given a graph H for which the near-k-twin relation of H is an equivalence of bounded index, produces a (labelled) graph G_H (on the same vertex set) of bounded degree, and a formula $\psi(x, y)$ such that $H = I_\psi(G_H)$.

The idea behind the procedure is the following: We start by dividing the near-k-twin classes of H into "small" and "large" ones (w.r.t. k), dealing with each of these two types of classes separately.

- Each large class (more precisely, the vertices in the class) is assigned a label and each pair of large classes receives another label indicating whether there are "almost all" or "almost none" edges between the two classes. The exceptions to "almost all" or "almost none" rules will be remembered by edges of the graph G_H (by Corollary 3.5 each vertex has a bounded number of such exceptions, hence the bounded degree of G_H). Using these labels and the graph G_H we properly encode the H-adjacency between the vertices in the large classes.
- The H-adjacency of the vertices from small equivalence classes (both within the small classes and also to the large ones) is encoded by assigning a new label to each such vertex and another new label to its neighbourhood. The vertices from small classes have no edges in the graph G_H.

Note that the construction sketched above depends on k and also on the number of near-k-twin equivalence classes of H. Unfortunately, as explained earlier, we cannot fix one universal value of the parameter k beforehand, but at least we can use upper bounds on both k and the number of equivalence classes (as in Definition 3.2). With a slightly more complicated use of labels, we can then give a universal formula $\psi(x, y)$ which depends only on the parameters k_0 and p of a (k_0, p)-near-uniform graph class $\mathcal{D}$, but is independent from particular $H \in \mathcal{D}$. This way we get a result even stronger than what is required for the proof of Theorem 3.3 (see Section 4 for more discussion):

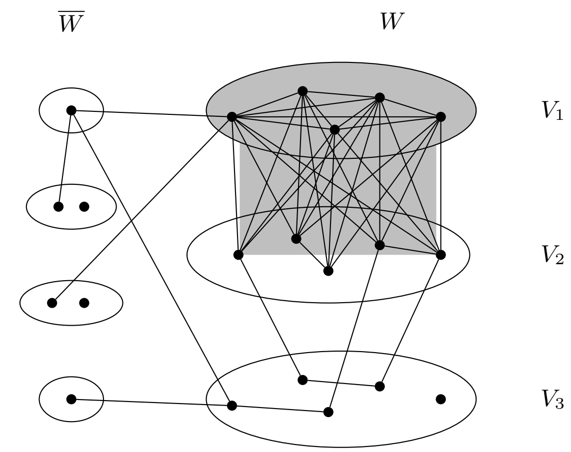

Figure 3. An illustration; small (on the left, $\overline{W}$) and large (on the right, W) near-k-twin classes of a graph H, and prevailing adjacencies within the large classes remembered by sets $F_1 = \{1\}$ and $F_2 = \{\{1, 2\}\}$, as in the proof of Theorem 3.7.

Theorem 3.7. *Let $k_0, p \in \mathbb{N}$, and $\mathcal{D}$ be a (k_0, p)-near-uniform graph class. There exists an FO formula $\psi(x, y)$, depending only on k_0 and p, such that $\mathcal{D} \subseteq I_\psi(\mathcal{G}_{2k_0p})$ where $\mathcal{G}_d$ denotes the class of (finite) graphs of degree at most d.*

Furthermore, for any $H \in \mathcal{D}$ and $k \leq k_0$ such that the near-k-twin relation of H is an equivalence of index at most p, one can in polynomial time compute a graph $G_H \in \mathcal{G}_{2k_0p}$ such that $H = I_\psi(G_H)$.

Proof. We are going to prove the theorem by defining the formula $\psi(x, y)$ and, for each $H \in \mathcal{D}$, efficiently constructing a graph $G_H \in \mathcal{G}_{2k_0p}$ such that $H = I_\psi(G_H)$. We give the construction of the graph G_H first, while postponing the definition of ψ to the end of the proof.

Let $0 \leq k \leq k_0$ be such that the near-k-twin relation of H is an equivalence of index at most p. Let $V_1, \ldots, V_m$ where $m \leq p$ be the near-k-twin classes of H with more than $5k$ vertices (possible "small" near-k-twin classes are ignored now). Observe that $W = V_1 \cup \cdots \cup V_m$ contains all but at most $5k(p - m) \leq 5k_0p$ vertices of H. Let $\overline{W} = V(H) \setminus W$ denote the remaining vertices in "small" equivalence classes. See an illustration in Figure 3.

We will construct the graph G_H in three stages. First, we define the graph $G_1 = (W, E_1 \cup E_2)$ on the set W, where the edge sets are given as:

- Let F_1 be the set of those indices i from $\{1, \ldots, m\}$ such that every vertex of V_i has at least $|V_i| - 2k$ neighbours in V_i (case (b) of Remark 3.6). We put $E_1 = \big\{\{u, v\} \mid u \neq v \wedge \exists i \in F_1 \text{ s.t. } u, v \in V_i\big\}$.
- Let F_2 be the set of those index pairs $\{i, j\}$ from $\{1, \ldots, m\}$ such that every vertex of V_i is connected to *all but* at most $2k$ vertices of V_j and every vertex of V_j is connected to *all but* at most $2k$ vertices of V_i (case (b) of Corollary 3.5). We put $E_2 = \big\{\{u, v\} \mid \exists \{i, j\} \in F_2 \text{ s.t. } u \in V_i \wedge v \in V_j\big\}$.

In the second step, we adjust G_1 by the original edges from H: Let $E_W = \{\{u, v\} \in E(H) \mid u, v \in W\}$. Then we put $G_2 = (W, E(G_1) \triangle E_W)$. See in Figure 4. Note that every vertex of G_2 has degree at most $2km$ by Corollary 3.5.

In the degenerate case of $k = 0$ we arrive at the same conclusion by the following alternative argument. By the definition, each near-

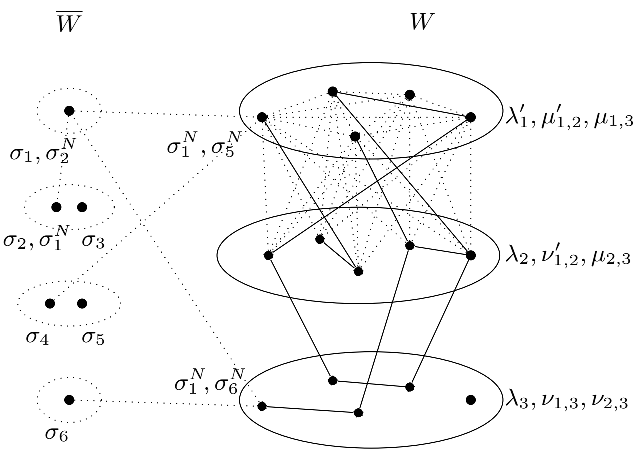

Figure 4. An illustration; graph G_2 of maximum degree 3 constructed for H (the dotted edges) from Figure 3, and the resulting labelling of $V(G_2) = W \cup \overline{W}$, as in the proof of Theorem 3.7.

0-twin class is an independent set and each pair of classes is again independent or induces a complete bipartite subgraph—this now defines G_1 and G_2 which is actually edgeless.

In the third step we add back the vertices from $\overline{W}$ (remember that $V(H) = W \cup \overline{W}$) by putting $G_H = (W \cup \overline{W}, E(G_2))$. Note that $G_H \in \mathcal{G}_{2km} \subseteq \mathcal{G}_{2k_0p}$,

Finally we label the vertices of G_H by the following fixed label set, which is independent of particular $H \in \mathcal{D}$:

$$\begin{aligned} L \;\; := \;\; & \{\lambda_i, \lambda_i' : i = 1, \dots, p\} \\ & \cup \{\mu_{i,j}, \nu_{i,j},\; \mu_{i,j}', \nu_{i,j}' : 1 \le i < j \le p\} \\ & \cup \{\sigma_j, \sigma_j^N : j = 1, \dots, 5k_0p\} \end{aligned}$$

The vertices of G_H are labelled as follows (see again Figure 4):

- For $i = 1, \dots, m \le p$, each vertex of V_i is assigned label λ_i' if $i \in F_1$, and label λ_i otherwise.
- For $1 \le i < j \le m \le p$, each vertex of V_i is assigned label $\mu_{i,j}'$ and each of V_j label $\nu_{i,j}'$ if $\{i, j\} \in F_2$, and labels $\mu_{i,j}$ and $\nu_{i,j}$, respectively, if $\{i, j\} \notin F_2$.
- Let $\overline{W} = \{w_1, w_2, \dots, w_r\}$ be indexed in any chosen order. For $j = 1, \dots, r \le 5k_0p$, the vertex w_j is assigned label σ_j and each neighbour of w_j in H is assigned label σ_j^N.

With G_H in place, we can now define the formula

$$\psi(x, y) \;\equiv\; (x \neq y) \wedge (\psi'(x, y) \vee \psi'(y, x))$$

where

$$\begin{aligned} \psi'(x, y) \equiv & \bigvee_{1 \le i \le p} \big(\lambda_i(x) \wedge \lambda_i(y) \wedge edge(x, y)\big) \\ \vee & \bigvee_{1 \le i \le p} \big(\lambda_i'(x) \wedge \lambda_i'(y) \wedge \neg edge(x, y)\big) \\ \vee & \bigvee_{1 \le i < j \le p} \big(\mu_{i,j}(x) \wedge \nu_{i,j}(y) \wedge edge(x, y)\big) \\ \vee & \bigvee_{1 \le i < j \le p} \big(\mu_{i,j}'(x) \wedge \nu_{i,j}'(y) \wedge \neg edge(x, y)\big) \\ \vee & \bigvee_{1 \le j \le 5k_0p} \big(\sigma_j(x) \wedge \sigma_j^N(y)\big). \end{aligned}$$

Clearly, $\psi(x, y)$ is independent of particular $H \in \mathcal{D}$ and depends only on the parameters k_0 and p. The construction of G_H from H and k is finished in polynomial time and it is also a simple routine to verify that $H = I_\psi(G_H)$. $\square$

This also finishes the proof of Theorem 3.3 via the fixed-parameter tractable algorithm of Seese [16].

4. Interpretability of graphs of bounded degree

Having defined near-uniform graph classes, and showing these classes can be FO interpreted in graph classes of bounded degree, it is a natural question to ask what is the exact relationship between those kinds of classes. As it turns out, we can prove (Theorem 4.3) that each class FO interpretable in a class of graphs of bounded degree is indeed near-uniform. Thus, near-uniform graph classes are exactly those graph classes, which are FO interpretable in graph classes of bounded degree.

4.1 Adjusted Gaifman's theorem

In the proof of the main result of this section we use the famous Gaifman's locality theorem [7] (see also [13]) about the local nature of the FO logic. However, for our purposes we need a specific variant of this theorem. To keep the paper self-contained, in this section we first recap the notation and statement of Gaifman's theorem and then state and prove a corollary tailored to our needs.

An FO formula $\phi(x_1, \dots, x_l)$ is *r-local*, sometimes denoted by $\phi^{(r)}(x_1, \dots, x_l)$, if for every graph G and all $v_1, \dots, v_l \in V(G)$ it holds $G \models \phi(v_1, \dots, v_l) \iff \bigcup_{1 \le i \le l} N_r^G(v_i) \models \phi(v_1, \dots, v_l)$, where $N_r^G(v)$ is the subgraph of G induced by v and all vertices of distance at most r from v.

Theorem 4.1 (Gaifman's locality theorem)**.** *Every first-order formula with free variables $x_1, \dots, x_l$ is equivalent to a Boolean combination of the following*

- *local formulas $\phi^{(r)}(x_1, \dots, x_l)$ around $x_1, \dots, x_l$, and*
- *basic local sentences, i.e. sentences of the form*

$$\exists x_1 \dots \exists x_k \left(\bigwedge_{1 \le i < j \le k} dist(x_i, x_j) > 2r \wedge \bigwedge_{1 \le i \le k} \phi^{(r)}(x_i) \right).$$

For a given q, the set of semantically different FO formulas ϕ of quantifier rank $qr(\phi) \le q$ with one free variable is finite. Clearly, this also holds for local FO formulas, as they are a special case of FO formulas. For a vertex v of a graph G, we define its local logical FO (q, r)-type as $\mathrm{tp}_{q,r}^G(v) = \{\phi^{(r)}(x) \mid G \models \phi^{(r)}(v) \text{ and } qr(\phi) \le q\}$.

It can be derived from Gaifman's theorem that if two vertices u and v are far apart in the graph, then whether $\psi(u, v)$ holds true depends only on the logical (q, r)-type of u and v, where q and r depend on ψ. This finding is formalized by the following (folklore) corollary of Theorem 4.1; as we were not able to find this precise formulation in the literature, we also provide a proof, for the sake of completeness.

Corollary 4.2. *For every FO formula $\psi(y, z)$ of two free variables there exist integers r and q such that the following holds true for any graph G: If $u, v_1, v_2 \in V(G)$ such that $dist(u, v_1) > 2r$, $dist(u, v_2) > 2r$ and $\mathrm{tp}_{q,r}^G(v_1) = \mathrm{tp}_{q,r}^G(v_2)$, then $G \models \psi(u, v_1)$ if and only if $G \models \psi(u, v_2)$.*

Proof. Let $\psi(y, z)$ be a formula, G a graph and $u, v_1, v_2 \in V(G)$ as in the statement of the Corollary. By Theorem 4.1, $\psi(y, z)$ is equivalent to a Boolean combination of local formulas $\phi^{(r)}(y, z)$ around y and z and basic local sentences. The validity of $\psi(y, z)$

for any choice of y and z therefore depends only on local formulas $\phi^{(r)}(y,z)$ around y and z (because the validity of basic local sentences is independent of the choice of y and z). Thus, whether $G \models \psi(u,v_1)$ holds true depends only on formulas $\phi^{(r)}(u,v_1)$ evaluated on the graph induced by $N_r^G(u) \cup N_r^G(v_1)$.

Because $\mathrm{dist}(u,v_1) > 2r$, this graph is actually a disjoint union of the graphs induced by $N_r^G(u)$ and $N_r^G(v_1)$. By the standard Ehrenfeucht-Fraisse games argument, validity of $N_r^G(u) \cup N_r^G(v_1) \models \phi^{(r)}(u,v_1)$ is then fully determined by the types $\mathrm{tp}^G_{q,r}(u)$ and $\mathrm{tp}^G_{q,r}(v_1)$ of u and v_1 respectively. The same reasoning can be applied to u and v_2, and since $\mathrm{tp}^G_{q,r}(v_1) = \mathrm{tp}^G_{q,r}(v_2)$, the result follows. □

4.2 Characterization of the interpretation

The following theorem provides us with a strong characterization of the classes FO-interpreted in graphs of degree at most d, in terms of near-k-twin equivalence. It amounts to, in an essence, the "opposite direction" to Theorem 3.7.

Theorem 4.3. *Let $\mathcal{G}_d$ be the class of (finite) graphs with maximum degree at most d and let $\psi(x,y)$ be an FO formula with two free variables. Then there exist k_0 and p, depending on d and ψ, such that for every $H \in I_\psi(\mathcal{G}_d)$ there exists $k \le k_0$ for which the near-k-twin relation of H is an equivalence of index at most p.*

Note that, for different graphs H, we may need different values of k (in particular, there may not be a universal value of k which would work for the whole class $I_\psi(\mathcal{G}_d)$).

Proof. Let $G \in \mathcal{G}_d$ such that $H = I_\psi(G)$. Recall that $V(H) = V(G)$ and $\{u,v\} \in E(H)$ if and only if $G \models \psi(u,v)$. Fixing G and H, we say that a vertex $x \in V(H)$ is *a-far from* $y \in V(H)$ if the graph distance from x to y in G is greater than a. From Corollary 4.2 and the fact that there exist altogether finitely many possible logical q-types of graph vertices (for each q), we immediately get that there exist integers c (take $c = 2r$ from Corollary 4.2) and m_0 depending on ψ such that the following claim holds true.

I. Every graph $H = I_\psi(G)$, for $G \in \mathcal{G}_d$, has a vertex partition $U_1 \cup \dots \cup U_m = V(H)$, where $m \le m_0$, satisfying the following: if $u, v \in U_i$ for some $i \in \{1,\dots,m\}$ and $X \subseteq V(G)$ denotes the set of vertices which are c-far from both u and v, then $N^H(u) \cap X = N^H(v) \cap X$. In other words, the neighbourhoods of u and v in H may differ only in vertices which are at a distance $\le c$ in G.

Since the maximum degree of G is at most d, any vertex of G has at most $1 + d(d-1)^{c-1} \le 1 + d^c$ vertices at distance up to c from it. Let $t_0 = 0$ and $t_i = |U_i|$ for $i = 1,\dots,m$.

Let us first briefly consider the case that $t_i \le t_0 + \dots + t_{i-1} + 4(1+d^c)$ for $i = 1,\dots,m$. Then $|V(H)| = t_0 + t_1 + \dots + t_m$ is bounded by a function of d and ψ, and so we can set $k := |V(H)|$ and thus have only one near-k-twin class of H.

So, for the rest of the proof, we assume that there exists $j \in \{1,\dots,m\}$ such that $t_j > t_0 + t_1 + \dots + t_{j-1} + 4(1+d^c)$. In this case we fix the least such index j and set $k := t_1 + \dots + t_{j-1} + 2(1+d^c)$. Note that k is again bounded by a function of d and ψ only. For any pair $u, v \in U_i$ where $i \in \{1,\dots,m\}$, u and v are near-k-twins since their neighbourhoods differ in at most $2(1+d^c)$ vertices by (I). For u, v from different parts of $(U_1,\dots,U_m)$, we say that a vertex w *distinguishes* u *from* v if w is c-far from both u, v and exactly one of wu, wv is an edge of H. If no vertex from $U_j \cup U_{j+1} \cup \dots \cup U_m$ distinguishes u from v, then again, u and v are in the near-k-twin relation of H since their neighbourhoods differ in at most $t_1 + \dots + t_{j-1} + 2(1+d^c) = k$ vertices.

On the other hand, assume that there exist $i \ne i' \in \{1,\dots,m\}$ such that some pair $u_0 \in U_i$, $v_0 \in U_{i'}$ is distinguished by a vertex $w_0 \in U_\ell$ where $j \le \ell \le m$. Up to symmetry, $u_0w_0 \in E(H)$ while $v_0w_0 \notin E(H)$. See an illustration in Figure 5. Consider any pair $u \in U_i$, $v \in U_{i'}$. Since $t_\ell \ge t_j > 2(1+d^c)$, there exists a vertex $w_1 \in U_\ell$ which is c-far from both u_0, u. Since $u_0w_0 \in E(H)$, we get $u_0w_1 \in E(H)$ and then $uw_1 \in E(H)$ by (I). Hence there are at least $t_\ell - (1+d^c)$ vertices in U_ℓ which are c-far from u and are neighbours of u. By symmetry, there exist at least $t_\ell - (1+d^c)$ vertices in U_ℓ which are c-far from v and are not neighbours of v. Altogether, we have got at least $t_\ell - 2(1+d^c) \ge t_j - 2(1+d^c) > k$ vertices in $N^H(u) \,\triangle\, N^H(v)$ and so u, v are not in the near-k-twin relation of H.

To recapitulate, u and v are near-k-twins if and only if both u, v are from the same part of $(U_1,\dots,U_m)$, or u, v are from distinct parts $U_i, U_{i'}$ such that no pair of representatives of $U_i, U_{i'}$ can be distinguished by a vertex from $U_j \cup U_{j+1} \cup \dots \cup U_m$. The near-$k$-twin relation of H is thus a coarsening of the partition $(U_1,\dots,U_m)$, and so it is an equivalence of index at most m. We can now choose $p := m$ and k_0 to be the maximum of the possible (and bounded) values of k considered above. □

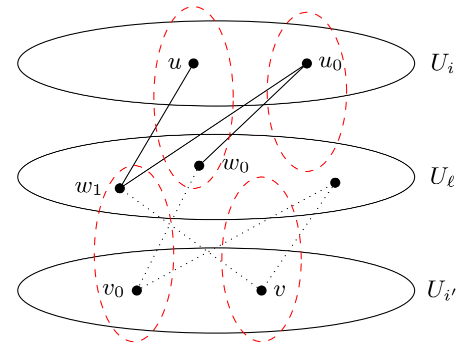

Figure 5. An illustration; there are (many) vertices w_1 in U_ℓ which are neighbours of $u \in U_i$ and at the same time not neighbours of $v \in U_{i'}$, as argued in the proof of Theorem 4.3. The dotted lines in the picture represent non-edges, i.e. nonadjacent pairs of vertices. The dashed (red) ellipses mark c-neighbourhoods of the respective vertices u, u_0, v_0, v.

Putting together the results of Theorems 3.7 and 4.3, we easily get also the following corollary which is interesting on its own:

Corollary 4.4. *Let $\mathcal{D}$ be a near-uniform graph class, and $\sigma(x,y)$ be an FO formula with two free variables. Then the class $I_\sigma(\mathcal{D})$ is again a near-uniform graph class.*

Proof. By Theorem 3.7, there exists an FO formula $\psi(x,y)$ such that $\mathcal{D} \subseteq I_\psi(\mathcal{G}_d)$ for suitable degree bound d depending on $\mathcal{D}$. We construct a formula $\sigma'(x,y)$ from $\sigma(x,y)$ by replacing every occurrence of $edge(z,z')$ with $\psi(z,z')$. Then $\mathcal{D}' = I_\sigma(\mathcal{D}) \subseteq I_{\sigma'}(\mathcal{G}_d)$. By Theorem 4.3, there exist k_0 and p such that, for every $H \in \mathcal{D}'$, there exists $k \le k_0$ for which the near-k-twin relation of H is an equivalence of index at most p. Consequently, $\mathcal{D}'$ is a near-uniform graph class. □

4.3 Characterization of ordinary interpretation

We now briefly return to ordinary meaning of an interpretation, in which one is allowed to select, in addition to new edges, also a subset of the domain. We easily argue that if we are given a graph H

obtained as an interpretation of some graph $G \in \mathcal{G}_d$ using formulas $\nu(x)$ and $\mu(x, y)$ so that $V(H) = \{u \in V(G) \mid G \models \nu(u)\}$ and $E(H) = \{\{u, v\} \subseteq V(H) \mid G \models \mu(u, v)\}$ then, again, there exist k_0, p independent of particular G such that the near-k-twin relation of H is an equivalence with at most p classes for some $k \leq k_0$.

We proceed as follows:

i. We consider the formula $\psi(x, y) \equiv \nu(x) \wedge \nu(y) \wedge \mu(x, y)$ and the graph $H' = I_\psi(G)$. Notice that $V(H) \subseteq V(H')$ and the vertices from $V(H') \setminus V(H)$ are isolated (they have no incident edges) by the definition of $\psi(x, y)$.

ii. We apply Theorem 4.3 to H' and $\psi(x, y)$ to obtain k_0 and p such that the near-k-twin relation of H' is an equivalence with at most p classes on $V(H')$ for some $k \leq k_0$.

iii. Since deleting an isolated vertex does not change the near-k-twin relation on the remaining vertices, after deleting the vertices $V(H') \setminus V(H)$ from H' to obtain H, the near-k-twin relation of H is still an equivalence with at most p classes.

4.4 Hardness of recognizing an interpretation

Recall the aforementioned result [14] claiming that it is NP-hard to decide whether a given graph is a *square* of some graph. The square of a graph can be straightforwardly described by an FO interpretation with $\psi_s(x, y) \equiv edge(x, y) \vee [x \neq y \wedge \exists z(edge(x, z) \wedge edge(z, y))]$, expressing that edges of the square are original edges or pairs at distance exactly two.

In our context, [14] hence means that there exist a graph class $\mathcal{C}$ and an FO formula $\psi(x, y)$ such that the problem, for a given graph $H \in I_\psi(\mathcal{C})$, to find $G \in \mathcal{C}$ such that $H = I_\psi(G)$ is not efficiently solvable (unless P=NP). Though, the reduction of [14] requires a class $\mathcal{C}$ of unbounded maximum degree while we are primarily interested in interpretations of the classes $\mathcal{G}_d$ of graphs of degrees at most d. Here we show a simple alternative reduction working already with the class $\mathcal{G}_3$ of graphs of degree at most 3.

Notice that such a result is not in a contradiction with Theorem 3.7 since each of the two results speaks about a different particular formula(s) ψ.

Theorem 4.5. *Let $\mathcal{G}_3$ denote the class of graphs of degree at most 3. There exists an FO formula $\psi_0(x, y)$ such that the problem, for a given graph $H \in I_{\psi_0}(\mathcal{G}_3)$, to find $G \in \mathcal{G}_3$ such that $H = I_{\psi_0}(G)$ is NP-hard.*

Proof. We reduce from the folklore NP-hard problem of 3-colouring a given 4-regular graph H_0. We construct a graph H from an arbitrary 4-regular graph H_0 as follows:

- Every vertex v of H_0 is replaced with a graph T_v which is a copy of the graph in Figure 6 including the dashed edges.
- Every edge e of H_0 is replaced with a graph U_e which is a copy of the graph in Figure 7 including the dashed edges.
- For every edge $e = \{u, v\}$ of H_0, the terminal e^1 of U_e is identified with u^i of T_u, and e^2 of U_e is identified with v^j of T_v, where e is the i-th edge at u and the j-th edge at v (for arbitrarily chosen orderings of edges incident to u, v).

The construction of H is independent of whether H_0 is 3-colourable. Note that since U_e contains a vertex of degree 5, it is $H \notin \mathcal{G}_3$.

Before defining the formula ψ_0, we briefly explain the underlying idea of the reduction. For a suitable subgraph G of H (on the same vertex set), we would like to have $H = I_{\psi_0}(G)$ if and only if every vertex gadget (of a vertex of H_0) restricted to G encodes one of three available colours (for this vertex in H_0), and every edge gadget in G "verifies" that the ends of the edge (in H_0) receive distinct colours.

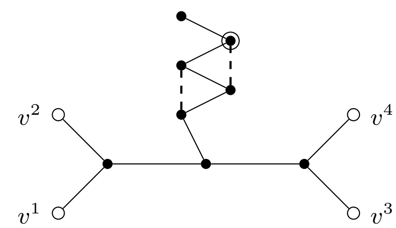

Figure 6. The vertex gadget T_v in the proof of Theorem 4.5.

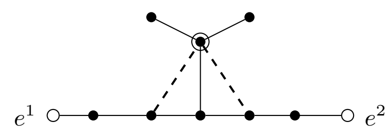

Figure 7. The edge gadget U_e in the proof of Theorem 4.5.

The above rough sketch is made precise now. Considering colours $1, 2, 3$, we define three *reduced vertex gadgets* of a vertex $v \in V(H_0)$ as $T_v^1 = T_v$ and T_v^2, T_v^3 obtained from T_v by removing one or the other dashed edge of T_v in Figure 6. Similarly, a *reduced edge gadget* U'_e of an edge $e \in E(H_0)$ is obtained from U_e in Figure 7 by removing both dashed edges. Assuming any 3-colouring $c : V(H_0) \to \{1, 2, 3\}$, we construct a graph $G \in \mathcal{D}_3$ analogously to the above construction of H, while replacing every vertex $v \in V(H_0)$ with $T_v^{c(v)}$ and every edge $e \in E(H_0)$ with U'_e.

Note that $G \subset H$. We call a vertex w a *v-marker* if w is adjacent to precisely one vertex of degree 1, and we call w an *e-marker* if w is adjacent to two vertices of degree 1 (see the circled vertices in Figures 6 and 7, respectively). Then every e-marker w of G belongs to some U'_e of $e = \{u, v\} \in E(H_0)$, and there are precisely two v-markers of G at distance 9 from w belonging to T_u^i and to T_v^j. We would now like to "verify" that the colouring c is proper, i.e. that $i \neq j$, in the formula ψ_0.

We define $\psi_0(x, y) \equiv edge(x, y) \vee \nu(x, y) \vee \eta(x, y)$ where

- $\nu(x, y)$ asserts that there exists z which is a neighbour of x or y, such that z is a v-marker and the 5-neighbourhood of z is isomorphic to one of T_v^1, T_v^2, T_v^3, and that x, y are the ends of one of the dashed edges in Figure 6;
- $\eta(x, y)$ asserts that one of x, y, say x, is an e-marker, y is at distance two from x, and the following holds: there exist vertices z, z' at distance 9 from x such that z, z' are v-markers with their 5-neighbourhoods isomorphic to T_v^i and T_v^j where $i \neq j$.

It is routine to rewrite the above description into an FO formula.

Clearly, $H = I_{\psi_0}(G)$ if and only if the above colouring c is proper. Conversely, it remains to prove that if $H = I_{\psi_0}(G)$ for any $G \in \mathcal{G}_3$, then H_0 is 3-colourable. Notice that $G \subseteq H$ and that the formula ψ_0 does not "add" edges to degree-1 vertices, and so the degree-1 vertices of G must be in a one-to-one correspondence with the v-marker and e-marker vertices of H.

Fix an e-marker w belonging to $U_e \subseteq H$. Since w is of degree 5 in H and of degree ≤ 3 in $G \in \mathcal{D}_3$, it is $G \models \psi_0(w, t)$ for some (actually, at least two) neighbour t of w in H. In particular, by the definition of $\eta(w, t)$, this means there exist two v-markers w', w'' at distance 9 from w in G. From the construction of H we know that w', w'' belong to T_u, T_v, respectively, where u, v are the ends of e in H_0. Again by $G \models \psi_0(w, t)$, the subgraph of G induced by $V(T_u)$ is one of T_u^1, T_u^2, T_u^3, say it is T_u^i. Similarly, the subgraph of G induced by $V(T_v)$ is, say, T_v^j and $i \neq j$. Since the same holds for any edge of H_0, an (arbitrary) graph $G \in \mathcal{D}_3$ such that $H = I_{\psi_0}(G)$ indeed encodes a proper 3-colouring of H_0. $\square$

5. Questions and open problems

Our approach and results open several natural questions which we believe are worth further investigation. Namely:

1. Can one characterize under which conditions on a formula $\psi(x, y)$ and a graph class $\mathcal{C}$, the following holds? Given a graph $H \in \mathcal{D}$ as an input, it would be possible to compute in polynomial (or in FPT with respect to ψ and $\mathcal{C}$) time a graph $G \in \mathcal{C}$ such that $H = \psi(G)$. Compare this to Theorems 3.7 and 4.5.
2. It is easy to generalize the notion of near-k-twins u, v in such a way that it would measure not the size of the symmetric difference between the neighbourhoods, $|N(u) \triangle N(v)|$, but structural properties of the subgraph induced on $N(u) \triangle N(v)$. For example, we may define a *near-sd_k-twin relation*, in which two vertices u, v would be near-sd_k-twins if the subgraph induced on $N(u) \triangle N(v)$ has shrub-depth at most k (see [9] for the definition of shrub-depth). One may then consider graph classes where the near-sd_k-twin relation is an equivalence. Is there an FPT algorithm for FO model checking on such graph classes?
3. Is it possible to extend our results to graph classes interpretable in more general sparse graph classes? For example, what is a characterization of graph classes interpretable in planar graphs? In graph classes of bounded expansion? Are there FPT algorithms for FO model checking on such classes?
4. In relation to the previous point, we know from Corollary 4.4 that the notion of near-uniform graph classes is robust under FO interpretations. We know of (at least) two other examples of such behaviour – the graph classes of bounded clique-width [2] and the graph classes of bounded shrub-depth [9] (which are robust even under MSO interpretations). Can one come up with other natural and interesting graph properties defining graph classes robust under FO interpretations?
5. Inspired by the classification of sparse graph classes by Nešetřil and Ossona de Mendez [15], we may investigate graph classes $\mathcal{D}$ with the property that, for every FO formula $\psi(x, y)$ there exists a graph F_ψ (as "forbidden") such that F_ψ is not present as an induced subgraph in any member of $I_\psi(\mathcal{D})$. This logical definition may be considered in analogy to the structural definition(s) of nowhere dense classes [15]. Can we say that such a class $\mathcal{D}$ is FO interpretable in some nowhere dense class?

To conclude, we make the following two explicit conjectures related to points 3 and 5 of the discussion.

Conjecture 5.1. *Let $\mathcal{C}$ be a nowhere dense graph class and $\mathcal{D}$ a graph class FO interpretable in $\mathcal{C}$. Then $\mathcal{D}$ has an FPT algorithm for FO model checking.*

Conjecture 5.2. *Let $\mathcal{D}$ be a graph class with the following property: for every FO formula $\psi(x, y)$ there exists a graph F_ψ such that F_ψ is not an induced subgraph of any member of $I_\psi(\mathcal{D})$. Then the class $\mathcal{D}$ is FO interpretable in some nowhere dense graph class.*

Acknowledgments

We would like to thank the anonymous referees for careful reading of the paper and for the many helpful suggestions and remarks.

J. Gajarský, P. Hliněný and J. Obdržálek have been supported by the Czech Science Foundation, project no. 14-03501S. M. S. Ramanujan acknowledges support from the Austrian Science Fund (FWF), project P26696 X-TRACT.

References

[1] S. Bova, R. Ganian, and S. Szeider. Model checking existential logic on partially ordered sets. In *CSL-LICS'14*. ACM, 2014. Article No. 21.

[2] B. Courcelle and S. Olariu. Upper bounds to the clique width of graphs. *Discrete Appl. Math.*, 101(1-3):77–114, 2000.

[3] A. Dawar, M. Grohe, and S. Kreutzer. Locally excluding a minor. In *LICS'07*, pages 270–279. IEEE Computer Society, 2007.

[4] G. Dong, L. Libkin, and L. Wong. Local properties of query languages. In *Database Theory - ICDT '97, Proceedings*, volume 1186 of *Lecture Notes in Computer Science*, pages 140–154. Springer, 1997.

[5] Z. Dvořák, D. Kráľ, and R. Thomas. Deciding first-order properties for sparse graphs. In *FOCS'10*, pages 133–142. IEEE Computer Society, 2010.

[6] M. Frick and M. Grohe. Deciding first-order properties of locally tree-decomposable structures. *J. ACM*, 48(6):1184–1206, 2001.

[7] H. Gaifman. On local and non-local properties. In *Proceedings of the Herbrand Symposium*, volume 107 of *Stud. Logic Found. Math.*, pages 105 – 135. Elsevier, 1982.

[8] J. Gajarský, P. Hliněný, D. Lokshtanov, J. Obdržálek, S. Ordyniak, M. S. Ramanujan, and S. Saurabh. FO model checking on posets of bounded width. In *FOCS'15*, pages 963–974. IEEE Computer Society, 2015.

[9] R. Ganian, P. Hliněný, J. Nešetřil, J. Obdržálek, P. O. de Mendez, and R. Ramadurai. When trees grow low: Shrubs and fast MSO_1. In *MFCS'12*, volume 7464 of *LNCS*, pages 419–430. Springer, 2012.

[10] R. Ganian, P. Hliněný, D. Kráľ, J. Obdržálek, J. Schwartz, and J. Teska. FO model checking of interval graphs. In *ICALP 2013, Part II*, volume 7966 of *LNCS*, pages 250–262. Springer, 2013.

[11] M. Grohe, S. Kreutzer, and S. Siebertz. Deciding first-order properties of nowhere dense graphs. In *STOC'14*, pages 89–98. ACM, 2014.

[12] M. Lampis. Algorithmic meta-theorems for restrictions of treewidth. In *ESA'10*, pages 549–560, 2010.

[13] L. Libkin. *Elements of Finite Model Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2004.

[14] R. Motwani and M. Sudan. Computing roots of graphs is hard. *Discrete Applied Mathematics*, 54(1):81–88, 1994.

[15] J. Nešetřil and P. Ossona de Mendez. *Sparsity: Graphs, Structures, and Algorithms*, volume 28 of *Algorithms and Combinatorics*. Springer, 2012.

[16] D. Seese. Linear time computable problems and first-order descriptions. *Math. Structures Comput. Sci.*, 6(6):505–526, 1996.