Approximation Schemes for First-Order Definable Optimisation Problems

Anuj Dawar*, Martin Grohe**, Stephan Kreutzer**, Nicole Schweikardt**

* University of Cambridge, U.K., Email: anuj.dawar@cl.cam.ac.uk
** Humboldt-Universität zu Berlin, Germany, Email: {grohe | kreutzer | schweika}@informatik.hu-berlin.de

Abstract

Let $\varphi(X)$ be a first-order formula in the language of graphs that has a free set variable $X$, and assume that $X$ only occurs positively in $\varphi(X)$. Then a natural minimisation problem associated with $\varphi(X)$ is to find, in a given graph $G$, a vertex set $S$ of minimum size such that $G$ satisfies $\varphi(S)$. Similarly, if $X$ only occurs negatively in $\varphi(X)$, then $\varphi(X)$ defines a maximisation problem. Many well-known optimisation problems are first-order definable in this sense, for example, MINIMUM DOMINATING SET or MAXIMUM INDEPENDENT SET.

We prove that for each class $\mathcal{C}$ of graphs with excluded minors, in particular for each class of planar graphs, the restriction of a first-order definable optimisation problem to the class $\mathcal{C}$ has a polynomial time approximation scheme.

A crucial building block of the proof of this approximability result is a version of Gaifman’s locality theorem for formulas positive in a set variable. This result may be of independent interest.

1. Introduction

It has long been known that many combinatorial optimisation problems that are hard to approximate in general have polynomial time approximation schemes (PTASs) on planar graphs, that is, they can be approximated to any given ratio in polynomial time. Among these problems are MINIMUM DOMINATING SET, MINIMUM VERTEX COVER, and MAXIMUM INDEPENDENT SET. The two main techniques for proving such approximation results on planar graphs are Lipton and Tarjan’s planar separator theorem [20] and Baker’s layerwise decomposition technique [4]. Both techniques have been generalised from planar graphs to more general graph classes such as graphs of bounded genus and ultimately to arbitrary classes of graphs with excluded minors [1, 10, 14, 7]. Recall that a minor of a graph $G$ is a graph that can be obtained from a subgraph of $G$ by contracting edges. We say that a class $\mathcal{C}$ of graphs has an excluded minor if there is some graph $H$ such that $H$ is not a minor of any graph in $G$. For example, the complete graph on five vertices, $K_5$, is an excluded minor of the class of all planar graphs. Most approximation results on general classes of graphs with excluded minors make heavy use of Robertson and Seymour’s structure theory for graphs with excluded minors [23]. In a recent paper, Demaine, Hajiaghayi, and Kawarabayashi [7] have proved algorithmic versions of some of the central parts of this theory and use these to obtain several new approximability results.

What kind of problems are approximable on graphs with excluded minors? Demaine et al. [7] gave a general criterion that is met by most problems known to be approximable, but is somewhat unsatisfactory because it describes when a certain proof technique works rather than describing a “natural” class of problems. On planar graphs, Khanna and Motwani [17] tried a more systematic “syntactic” approach: They defined three “generic” problems based on propositional logic and showed that the planar versions of these problems have PTASs. Then they showed that most problems which at that time were known to have PTASs can easily be reduced to one of these three problems. In this paper, we carry out a different logic based approach towards identifying a large class of problems that have PTASs on classes of graphs with excluded minors. Our approach, in contrast to that of Khanna and Motwani, is based on first-order logic: Let $\varphi(X)$ be a first-order formula in the language of graphs that has a free set variable $X$, and assume that $X$ only occurs positively in $\varphi(X)$. Then a natural minimisation problem $\text{MIN}_{\varphi(X)}$ associated with $\varphi(X)$ is to find, in a given graph $G$, a vertex set $S$ of minimum size such that $G$ satisfies $\varphi(S)$. Many natural minimisation problems can be described as problems $\text{MIN}_{\varphi(X)}$ for a suitable formula $\varphi(X)$. For example, the MINIMUM DOMINATING SET problem is $\text{MIN}_{\varphi(X)}$ for the formula $\varphi(X) = \forall x (Xx \lor \exists y (Xy \land Exy))$. The condition that $\varphi(X)$ be positive in $X$ is imposed to guarantee monotonicity, which is necessary to exclude pathological examples (see Example 11). Similarly, if $X$ only occurs negatively in a formula $\psi(X)$ then this formula defines a natural maximisation problem $\text{MAX}_{\psi(X)}$. For example, MAX-
arbitrary relational structures and also to weighted versions above. Our theorem can easily be extended from graphs to order definable in our sense. Even on planar graphs, our included minors. Indeed, it is easy to find problems that meet restrictions of first-order definable optimisation problems to the full version of this paper. Even with these generalisations, we do not claim that our theorem captures all problems that have PTASs on classes of graphs with excluded minors. Indeed, it is easy to find problems that meet Demaine et al.’s approximability criterion, but are not first-order definable in our sense. Even on planar graphs, our approach seems incomparable with Khanna and Motwani’s in that there is no obvious translation of our syntactically defined problems into theirs or vice versa. An important difference between our result and those of Demaine et al. and Khanna and Motwani is that we obtain an EPTAS. For the problem PLANAR TMIN, for which Khanna and Motwani obtained a PTAS, it can actually be proved that, under reasonable complexity theoretic assumptions, it does not have an EPTAS [5, 21].

The proof of Theorem 1 has two parts: the second, algorithmic, part builds on techniques that were first applied in [14] to classes of graphs with excluded minors and generalise Baker’s layerwise decomposition technique [4]. However, the techniques have to be generalised considerably to handle the very general class of problems we consider here. The crucial property of first-order definable optimisation problems that our algorithms exploit is the locality of first-order logic. In the first part of the proof of Theorem 1, we prove a “positive version” of Gaifman’s locality theorem, a result which may be of independent interest:

**Theorem 2.** Let \( \exists \psi(X) \) be a first-order sentence that is positive in the set variable \( X \). Then there is a Boolean combination \( \psi(X) \) of basic local sentences so that \( \psi(X) \) is positive in \( X \) and equivalent to \( \psi(X) \).

The necessary definitions will be given later. Rather unexpectedly, the proof of this theorem proved to be fairly difficult, as we were unable to adapt the known proofs of Gaifman’s theorem [12] (see [9, 16] for alternative proofs) or of its existential version [15]. Our proof of the positive version uses ideas from [3] to analyse the spatial distribution of the types occurring in a structure, and it uses a lemma from [15] to get from a nonuniform to a uniform version of the theorem, but the core combinatorial argument is new.

### 2. Preliminaries

A **vocabulary** is a finite set of relation symbols and constant symbols. Associated with every relation symbol \( R \) is a positive integer called the **arity** of \( R \). In the following, \( \tau \) always denotes a vocabulary. \( \tau \) is called **relational** if it does not contain any constant symbol.

A **\( \tau \)-structure** \( \mathcal{A} \) consists of a non-empty set \( A \), called the **universe** of \( \mathcal{A} \), an element \( c^{\mathcal{A}} \in A \) for each constant symbol \( c \in \tau \), and a relation \( R^{\mathcal{A}} \subseteq A^r \) for each \( r \)-ary relation symbol \( R \in \tau \).

The **Gaifman graph** of a **\( \tau \)-structure** \( \mathcal{A} \) is the (undirected, loop-free) graph \( G_{\mathcal{A}} \) with vertex set \( A \) and an edge between two vertices \( a, b \in A \) iff there exists an \( R \in \tau \) and a tuple \( (a_1, \ldots, a_r) \in R^{\mathcal{A}} \) such that \( a, b \in \{ a_1, \ldots, a_r \} \).

The **distance** between two elements \( a, b \in A \) in \( \mathcal{A} \), denoted by \( \text{dist}^\mathcal{A}(a, b) \), is defined to be the length (that is, number of edges) of the shortest path from \( a \) to \( b \) in the Gaif-
man graph of \( \mathcal{A} \). For \( r \geq 0 \) and \( a \in A \), the \( r \)-neighbourhood of \( a \) in \( \mathcal{A} \) is the set \( N^{(a)}(a) = \{ b \in A : \text{dist}(a,b) \leq r \} \).

The induced substructure of \( \mathcal{A} \) with universe \( N^{(a)}(a) \) is denoted by \( \mathcal{A}^{(a)} \). We omit superscripts \(^{(a)}\) if \( \mathcal{A}^{(a)} \) is clear from the context.

We write \( \text{FO}(\tau) \) to denote the class of all formulae in first-order logic over the vocabulary \( \tau \), and we write \( q \varphi(\phi) \) to denote the quantifier rank of an \( \text{FO}(\tau) \)-formula \( \phi \). If \( \mathcal{X} \) is a unary relation symbol not in \( \tau \), then an occurrence of \( \mathcal{X} \) in an \( \text{FO}(\tau \cup \{ \mathcal{X} \}) \)-formula \( \phi \) is said to be positive if it is within the scope of an even number of negations and it is said to be negative otherwise. We say that the formula \( \phi \) is positive in \( \mathcal{X} \) (or \( X \)-positive) if all occurrences of \( \mathcal{X} \) in \( \phi \) are positive. Similarly, we say that \( \phi \) is negative in \( \mathcal{X} \) (or \( X \)-negative) if all occurrences of \( \mathcal{X} \) in \( \phi \) are negative.

For every \( r \geq 0 \), we let \( \text{dist}_{\leq r}(x,y) \) be an \( \text{FO}(\tau) \)-formula expressing that the distance between \( x \) and \( y \) is at most \( r \). We often write \( \text{dist}(x,y) \leq r \) instead of \( \text{dist}_{\leq r}(x,y) \) and \( \text{dist}(x,y) > r \) or \( \text{dist}_{\leq r}(x,y) \) instead of \( \text{dist}(x,y) > r \).

The \( r \)-relativisation of a formula \( \varphi(x) \) is the formula \( \varphi^r(x) \) obtained from \( \varphi \) by first renaming all bound variables so that they are different from \( x \) and then replacing each subformula of the form \( \exists y \psi \) by \( \exists y(\text{dist}(x,y) \leq r \land \psi) \) and each subformula of the form \( \forall y \psi \) by \( \forall y(\text{dist}(x,y) \leq r \rightarrow \psi) \). Clearly, the \( r \)-relativisation of every formula \( \varphi(x) \) is \( r \)-local, that is, for every \( \tau \)-structure \( \mathcal{A} \) and every \( a \in A \) we have \( \mathcal{A} \models \varphi(a) \iff \mathcal{A}^r \models \varphi^r(a) \). Note that we also have \( \mathcal{A} \models \varphi(a) \iff \mathcal{A}^r \models \varphi^r(a) \). A (symmetric) basic local sentence (with parameters \( k, r, q \)) is a sentence of the form

\[
\exists x_1 \cdots \exists x_k \left( \bigwedge_{1 \leq i < j \leq k} \text{dist}(x_i,x_j) > 2r \land \bigwedge_{1 \leq i \leq k} \psi(x_i) \right),
\]

where \( k, r \geq 1 \) and \( \psi(x) \) is \( r \)-local and of quantifier rank \( q \) (here, the adjective “symmetric” emphasises that the same formula \( \psi \) is used for each of the variables \( x_i \)).

Theorem 3 (Gaifman [12]). Every first-order sentence over a relational vocabulary is equivalent to a Boolean combination of basic local sentences.

3. A positive locality theorem

In this section we present a proof of the version of Gaifman’s theorem for formulae positive in a unary relation symbol, stated in Theorem 2. From now on, fix a relational vocabulary \( \tau \) and a unary relation symbol \( X \not\in \tau \). For proving Theorem 2 we adopt the approach of [15] of using asymmetric basic local formulae in an intermediate step.

An asymmetric basic local sentence with parameters \( k, \kappa, r, q \) is a sentence of the form

\[
\exists x_1 \cdots \exists x_k \left( \bigwedge_{1 \leq i < j \leq k} \text{dist}(x_i,x_j) > \kappa \cdot 2r \land \bigwedge_{i=1}^k \psi(x_i) \right),
\]

where \( \psi(x_i) \) is \( r \)-local and of quantifier rank at most \( q \). We denote the set of all asymmetric basic local sentences with parameters \( k' \leq k, \kappa, r' \leq r, \) and \( q' \leq q \) by \( \text{ABL}(k, \kappa, r, q) \). By \( \text{ABL}^+(k, \kappa, r, q) \) (respectively, \( \text{ABL}^-(k, \kappa, r, q) \)) we denote the set of all sentences in \( \text{ABL}(k, \kappa, r, q) \) that are positive (respectively, negative) in \( \mathcal{X} \).

Similarly, we write \( \text{BL}(k, r, q) \), \( \text{BL}^+(k, r, q) \), and \( \text{BL}^-(k, r, q) \) for, respectively, the set of all, all \( X \)-positive, and all \( X \)-negative asymmetric basic local sentences with parameters \( k' \leq k, r' \leq r, \) and \( q' \leq q \).

For a sentence \( \varphi \in \text{ABL}(k, \kappa, r, q) \) of the form

\[
\exists x_1 \cdots \exists x_k \left( \bigwedge_{1 \leq i < j \leq k} \text{dist}(x_i,x_j) > 2r \land \bigwedge_{i=1}^k \psi(x_i) \right)
\]

we write \( \varphi_{[1/k]} \) to denote the sentence

\[
\exists x_1 \cdots \exists x_k \left( \bigwedge_{1 \leq i < j \leq k} \text{dist}(x_i,x_j) > 2r \land \bigwedge_{i=1}^k \psi(x_i) \right)
\]

(in particular, \( \varphi_{[1/k]} \in \text{ABL}(k, 1, r, q) \)).

The two major steps in proving Theorem 2 consist of showing the following two technical lemmas:

**Lemma 4.** Let \( K, Q, R \geq 2 \) and let \( \kappa := 2^{\kappa^2-1} \). Suppose \( \mathcal{A} \) and \( \mathcal{B} \) are \( \tau \cup \{ X \} \)-structures such that every \( X \)-positive (resp. \( X \)-negative) basic local sentence in \( \text{BL}(K, \kappa, 2R, Q) \) that holds in \( \mathcal{A} \) also holds in \( \mathcal{B} \). Then we have for every \( X \)-positive (resp. \( X \)-negative) sentence \( \varphi \in \text{ABL}(K, \kappa, Q, R) \) that \( \mathcal{A} \models \varphi \) implies \( \mathcal{B} \models \varphi \).

We omit the proof of Lemma 4 since it is virtually identical to the proof of Lemma 4 in [15]. We will use Lemma 4 as an intermediate step in proving the following:

**Lemma 5.** For every \( q \geq 0 \) there exist \( K, R, Q \geq 2 \) such that for all \( \tau \cup \{ X \} \)-structures \( \mathcal{A} \), \( \mathcal{B} \) the following holds: If for every \( \varphi \in \text{BL}^{+}(K, R, Q) \), \( \mathcal{A} \models \varphi \) implies \( \mathcal{B} \models \varphi \), and for every \( \varphi \in \text{BL}^{-}(K, R, Q) \), \( \mathcal{B} \models \varphi \) implies \( \mathcal{A} \models \varphi \), then we have for every \( X \)-positive \( \text{FO}(\tau \cup \{ X \}) \)-sentence \( \zeta \) of quantifier rank at most \( q \) that \( \mathcal{A} \models \zeta \) implies \( \mathcal{B} \models \zeta \).

Note that by using Lemma 5 one easily obtains a proof of Theorem 2 (details of this will be given in the full version of the paper).

The remainder of Section 3 is devoted to the proof of Lemma 5. To prove Lemma 5, we use the following “\( X \)-positive” variant of the classical Ehrenfeucht-Fraissé game (EF-game, for short) for first-order logic.

3.1. The \( X \)-positive EF-game.

The rules of this game are the same as for the “classical” EF-game for first-order logic (cf., e.g. [9]), the winning condition, however, is slightly different. To be precise, the “\( X \)-positive” EF-game is defined as follows:
Let \( q \) be a positive integer. The \( q \)-round \( X \)-positive EF-game is played by two players, the spoiler and the duplicator, on two \( \tau \cup \{X\} \)-structures \( \mathcal{A} \) and \( \mathcal{B} \). The spoiler’s intention is to show a difference between the two structures, while the duplicator tries to make them look alike. There is a fixed number \( q \) of rounds. Each round \( i \in \{1, \ldots, q\} \) is played as follows: First, the spoiler chooses either an element \( a_i \) in \( A \) or an element \( b_i \) in \( B \). Next, the duplicator chooses an element in the other structure. I.e., she chooses an element \( b_i \) in \( B \) if the spoiler’s move was in \( A \), or an element \( a_i \) in \( A \) if the spoiler’s move was in \( B \). After \( q \) rounds the game ends with elements \( a_1, \ldots, a_q \) chosen in \( A \) and \( b_1, \ldots, b_q \) chosen in \( B \). The duplicator has won the game if the mapping \( f \) defined via \((a_i \mapsto b_i)\) for \( 1 \leq i \leq q \) is an \( X \)-positive partial isomorphism from \( \mathcal{A} \) to \( \mathcal{B} \), i.e.,

(i) for any tuple of elements \((v_1, \ldots, v_s)\) within the domain of \( f \) and any relation symbol \( R \in \tau \) of arity \( s \), we have \((v_1, \ldots, v_s) \in R^\mathcal{A} \iff (f(v_1), \ldots, f(v_s)) \in R^\mathcal{B}\), and

(ii) for any element \( v \) within the domain of \( f \) and for the particular unary relation symbol \( X \), we have \( v \in X^\mathcal{A} \iff f(v) \in X^\mathcal{B}\).

Otherwise, the spoiler has won the game. Since the game is finite, one of the two players must have a \emph{winning strategy}, i.e., he or she can always win the game no matter how the other player plays. We write \( \mathcal{A} \equiv_{\exists} X^\mathcal{B} \) to denote that the duplicator has a winning strategy in the \( q \)-round \( X \)-positive EF-game on \( \mathcal{A} \) and \( \mathcal{B} \). Note that the relation defined by \( \equiv_{\exists} X^\mathcal{B} \) on the class of all \( \tau \cup \{X\} \)-structures is reflexive and transitive, but not symmetric.

The fundamental use of the \( q \)-round \( X \)-positive EF-game comes from the fact that it characterises definability by \( X \)-positive first-order sentences in the following sense:

**Proposition 6.** Let \( \mathcal{A} \) and \( \mathcal{B} \) be \( \tau \cup \{X\} \)-structures and let \( q \) be a positive integer. If \( \mathcal{A} \equiv_{\exists} X^\mathcal{B} \), then we have for every \( X \)-positive FO(\( \tau \cup \{X\} \))-sentence \( \zeta \) of quantifier rank at most \( q \), that \( \mathcal{A} \models \zeta \) implies \( \mathcal{B} \models \zeta \).

The proof is straightforward.

Now, note that Lemma 5 is an immediate consequence of Proposition 6 and the following lemma.

**Lemma 7.** For every \( q \geq 0 \) there exist \( K, R, Q \geq 2 \) such that for all \( \tau \cup \{X\} \)-structures \( \mathcal{A}, \mathcal{B} \): If for every \( \phi \in \text{BL}^+ (K, R, Q) \), \( \mathcal{A} \models \phi \) implies \( \mathcal{B} \models \phi \), and for every \( \phi \in \text{BL}^- (K, R, Q) \), \( \mathcal{B} \models \phi \) implies \( \mathcal{A} \models \phi \), then \( \mathcal{A} \equiv_{\exists} X^\mathcal{B} \).

Subsection 3.2 below is devoted to the proof of Lemma 7.

3.2. Proof of Lemma 7.

Before describing the duplicator’s winning strategy we need some preparation.

Let \( r, q \geq 0 \), \( \mathcal{A} \) a \( \tau \cup \{X\} \)-structure and \( a \in A \). The \textbf{full \((r,q)\)-type of \( a \) in } \( \mathcal{A} \) is the set \textbf{full-\((r,q)\)-type}\( ^{\mathcal{A}}(a) \) := \{ \phi'(x) : \phi \in \text{FO}(\tau \cup \{X\}) \text{ positive in } X, \text{ qr}(\phi) \leq q, \mathcal{A} \models \phi'(a) \} \).

Note that there is a formula \( \theta_{(r,q,\mathcal{A},a)}(x) := \bigwedge_{\phi \in \text{full-\((r,q)\)-type}^{\mathcal{A}}(a)} \phi(x) \) defining an element’s full \((r,q)\)-type and for all \( \tau \cup \{X\} \)-structures \( \mathcal{B} \) and all \( b \in B \) we have \( \mathcal{B} \models \theta_{(r,q,\mathcal{A},a)}(b) \iff \text{full-\((r,q)\)-type}^{\mathcal{A}}(b) = \text{full-\((r,q)\)-type}^{\mathcal{B}}(a) \).

Due to lack of space, we defer the proof of Lemma 8 to the full version of the paper. For the proof of Lemma 7 we also need the notions of \emph{positive} and \emph{negative} types of an element \( a \) in a \( \tau \cup \{X\} \)-structure \( \mathcal{A} \). The \emph{positive \((r,q)\)-type} of \( a \) is the set

\[ \text{pos-\((r,q)\)-type}^{\mathcal{A}}(a) := \{ \phi'(x) : \phi \in \text{FO}(\tau \cup \{X\}) \text{ positive in } X, \text{ qr}(\phi) \leq q, \mathcal{A} \models \phi'(a) \}. \]
Similarly, the negative \((r,q)\)-type of \(a\) is the set
\[
\text{neg-}(r,q)\text{-type}^a(a) := \{ \varphi'(x) : \varphi \in FO(\tau \cup \{X\}) \text{ negative in } X, q\varphi(\varphi) \leq q, \mathcal{A} \models \varphi'(a) \}.
\]

Note that \(\text{pos-}(r,q)\text{-type}^a(a) \subseteq \text{full-}(r,q)\text{-type}^a(a)\) and \(\text{neg-}(r,q)\text{-type}^a(a) \subseteq \text{full-}(r,q)\text{-type}^a(a)\). The formula
\[
\theta^+(r,q,a)(x) := \bigwedge_{\varphi \in \text{pos-}(r,q)\text{-type}^a(a)} \varphi(x),
\]
defines the positive \((r,q)\)-type in the sense that for all \((\tau \cup \{X\})\)-structures \(\mathcal{B}\) and all \(b \in B\) with \(\mathcal{B} \models \theta^+(r,q,a)(b)\) we have \(\text{pos-}(r,q)\text{-type}^a(b) \supseteq \text{pos-}(r,q)\text{-type}^a(a)\). Analogously, one obtains a formula \(\theta^-(r,q,a)(x)\) that defines the negative \((r,q)\)-type of \(a\) in \(\mathcal{A}\). Note that the formula \(\theta^+(r,q,a)\) and \(\theta^-(r,q,a)\) are \(r\)-local and of quantifier rank at most \(\bar{q}\) (where \(\bar{q} \geq q\) only depends on \(q, r,\) and \(\tau \cup \{X\}\)). Furthermore, \(\theta^+(r,q,a)(x)\) is positive in \(\mathcal{A}\), whereas \(\theta^-(r,q,a)(x)\) is negative in \(\mathcal{A}\). In the following, we often identify the types with these formulae.

We denote the set of all positive and negative \((r,q)\)-types by \(\Theta^+_r\) and \(\Theta^-_r\), respectively. (A positive or negative) type \(\theta(x)\) is realised in a structure \(\mathcal{A}\) if there is an \(a \in A\) such that \(\mathcal{A} \models \theta(a)\). We call \(a\) a realisation of \(\theta\) in \(\mathcal{A}\).

**Proof of Lemma 7:**
We fix \(q \geq 0\) and let \(k := q, r := 3q,\) and \(Q := q+1\). Let \(\hat{K}, \hat{R}\) be chosen according to Lemma 8. Now let \(\mathcal{A}\) and \(\mathcal{B}\) be \(\tau \cup \{X\}\)-structures such that
\[
(\ast) \text{ for every } \varphi \in \text{BL}^+(\hat{K}, 2^{\hat{K}^2-1} \cdot 10\hat{R} Q), \mathcal{A} \models \varphi \text{ implies } \mathcal{B} \models \varphi, \text{ and for every } \varphi \in \text{BL}^-(\hat{K}, 2^{\hat{K}^2-1} \cdot 10\hat{R} Q), \mathcal{B} \models \varphi \text{ implies } \mathcal{A} \models \varphi.
\]
Before we can describe the duplicator’s winning strategy in the \(q\)-round \(X\)-positive EF-game on \(\mathcal{A}\) and \(\mathcal{B}\), we first need to explore the “playing fields” \(\mathcal{A}\) and \(\mathcal{B}\). To this end, we first apply Lemma 8 to \(\mathcal{A}\) and \(\mathcal{B}\) (with \(k, r, q\)) to obtain numbers \(K \leq \hat{K}, R \leq \hat{R}\) and sets \(C := C^{\mathcal{A}} \subseteq A\) and \(D := C^{\mathcal{B}} \subseteq B\) such that for \(k := 2^{\hat{K}^2-1}\) and for every \(\mathcal{D} \in \{\mathcal{A}, \mathcal{B}\}\) the properties (1)–(3) of Lemma 8 are satisfied.

Note that, since \(K \leq \hat{K}\) and \(R \leq \hat{R}\), \((\ast)\) in particular holds when replacing \(\hat{K}\) with \(K\) and \(\hat{R}\) with \(R\). Thus, by applying Lemma 4 (both the \(X\)-positive and the \(X\)-negative version, while interchanging the roles of \(\mathcal{A}\) and \(\mathcal{B}\) when applying the \(X\)-negative version), we obtain
\[
(\ast\ast) \text{ for every } \varphi \in \text{ABL}^+(K, \kappa, 5R, Q), \mathcal{A} \models \varphi \text{ implies } \mathcal{B} \models \varphi\mid_{1/k} \text{ and for every } \varphi \in \text{ABL}^-(K, \kappa, 5R, Q), \mathcal{B} \models \varphi \text{ implies } \mathcal{A} \models \varphi\mid_{1/k}.
\]

Let us proceed with fixing some more notation. We call a full \((r,q)\)-type \(\theta \in \Theta^+_r\) rare in \(\mathcal{A}\) (in \(\mathcal{B}\)) if it is \(R\)-covered by \(C\) (by \(D\), respectively); otherwise, we call \(\theta\) frequent.

A positive \((r,q)\)-type \(\pi \in \Theta^+_r\) is called saturated if there is a full type \(\theta \in \Theta^+_r\) that is frequent in \(\mathcal{B}\) such that \(\pi \subseteq \theta\). Similarly, a negative \((r,q)\)-type \(\nu \in \Theta^-_r\) is called saturated if there is a full type \(\theta \in \Theta^-_r\) that is frequent in \(\mathcal{A}\) such that \(\nu \subseteq \theta\).

We define a bipartite Graph \(\mathcal{G}\) on \(C \cup D\) by drawing an edge from \(c \in C\) to \(d \in D\) if there are \(c' \in N^2_c(d)\) and a \(d' \in N^2_d(c)\) such that the positive \((4R,q+1)\)-type of \(c'\) is contained in the positive \((4R,q+1)\)-type of \(d'\) and hence the negative \((4R,q+1)\)-type of \(d'\) is contained in the negative \((4R,q+1)\)-type of \(c'\).

We call an element \(c \in C\) special if there is an \(a \in N^2_c\) such that the positive \((r,q)\)-type of \(a\) is not saturated (i.e., every full \((r,q)\)-type \(\theta \subseteq \text{pos-}(r,q)\text{-type}^a(a)\) is rare in \(\mathcal{B}\), i.e., realised only by elements in \(N\times(D)\)). Similarly, an element \(d \in D\) is called special if there is a \(b \in N^2_d\) such that the negative \((r,q)\)-type of \(b\) is not saturated.

Let \(C_S \subseteq C\) and \(D_S \subseteq D\) be the sets of all special vertices.

**Claim 9:** The bipartite graph \(\mathcal{G}\) has a matching \(M\) such that each special element of \(C\) and \(D\) is an endpoint of an edge in \(M\).

**Proof of Claim 9:** Let \(\ell = |C_S|\) and \(C_S = \{c_1, \ldots, c_\ell\}\). For every \(i \in \{1, \ldots, \ell\}\) let \(\pi_i(x)\) be the positive \((4R,q+1)\)-type of \(c_i\). Then \(\mathcal{A}\) satisfies the sentence \(\varphi_1\mid_{1/k}\) (to see this, recall the assumption \((\ast\ast)\) on \(\mathcal{A}\) and \(\mathcal{B}\) on the sentences in \(\text{ABL}^+(K, \kappa, 5R, Q)\), note that \(\ell \leq K\) since \(C_S \subseteq C\) and \(|C| \leq K\), and recall from Lemma 8 that elements in \(C\) have pairwise distance \(\geq 10R\)). Hence we can find \(b_1, \ldots, b_{\ell} \in B\) of pairwise distance greater than \(10R\) such that \(\mathcal{B} \models \pi_i(b_i)\) for every \(i \in \{1, \ldots, \ell\}\).

Let us next note that each of the \(b_i\) belongs to \(N^2\). This can be seen as follows. Since \(c_i\) is special, there exists an \(a_i \in N^2(c_i)\) whose positive \((r,q)\)-type \(\pi_i\) is not saturated, i.e., each full \((r,q)\)-type that contains \(\pi_i\) is realised in \(\mathcal{B}\) only by elements in \(N^2\). Since the positive \((4R,q+1)\)-type of \(b_i\) contains the positive \((4R,q+1)\)-type \(\pi_i\) of \(c_i\) and \(\mathcal{B}^\pi_i\) satisfies the \(X\)-positive formula \(\exists y(\text{dist}(c_i,y) \leq R \wedge \pi_i(y))\), we know that also \(\mathcal{B}^\pi_i\) satisfies this formula, and thus there exists an element \(b'_i\) with \(\text{dist}(b_i, b'_i) \leq R\) whose full \((r,q)\)-type contains \(\pi_i\). Since \(\pi_i\) is not saturated, we conclude that \(b'_i \in N^2\) and hence \(b_i \in N^2\).
this, note that $c'_i := c_i \in N_{2R}(c_i)$ has the positive $(4R, q+1)$-type $\pi_i$, and $d'_i := b_i \in N_{2R}(d_i)$ has a positive $(4R, q+1)$-type that contains $\pi_i$. It follows that each set $C' \subseteq C_S$ of special vertices has at least $|C'|$ neighbours in $D$.

Analogously, we can show that each set $D' \subseteq D_S$ of special vertices has at least $|D'|$ neighbours in $C$.

Now Claim 9 immediately follows from the following purely combinatorial lemma, which may be viewed as an extension of Hall’s well known marriage theorem. Let us say that a vertex is covered by a matching if it is an endpoint of an edge in the matching.

**Lemma 10.** Let $\mathcal{G}$ be a bipartite graph with bipartition $C, D$ of the vertex set. Let $C_S \subseteq C$ and $D_S \subseteq D$, and suppose that each $C' \subseteq C_S$ has at least $|C'|$ neighbours in $D$ and each $D' \subseteq D_S$ has at least $|D'|$ neighbours in $C$. Then there is a matching $M$ of $\mathcal{G}$ that covers each vertex in $C_S \cup D_S$.

The proof of Lemma 10 can be found in the full version of this paper. To proceed with the proof of Lemma 7 let us now fix a matching $M$ that covers all special vertices (such a matching exists by Claim 9). Let $c_1, \ldots, c_m \in C$ and $d_1, \ldots, d_m \in D$ be the vertices covered by $M$ via an edge between $c_i$ and $d_i$ for each $i \in \{1, \ldots, m\}$. By the definition of the graph $\mathcal{G}$, for $i \in \{1, \ldots, m\}$, there are vertices $c'_i \in N_{2R}(c_i)$ and $d'_i \in N_{2R}(d_i)$ such that the positive $(4R, q+1)$-type of $d'_i$ contains the positive $(4R, q+1)$-type of $c'_i$. In particular, the duplicator has a winning strategy for the $q$-round X-positive EF-game on $\mathcal{N}_{4R}(c'_i)$ and $\mathcal{N}_{4R}(d'_i)$.

Recall that, by the definition of special vertices, every $a \in A$ whose positive $(r, q)$-type is not saturated is in the $R$-neighbourhood of some special vertex of $C$ and hence, in particular, in $N_R(c_i) \subseteq N_{2R}(c'_i)$, for some $i \in \{1, \ldots, m\}$ (to see this, note that (1) $a$ has to belong to $N_R(c)$ due to Lemma 8, and (2) the vertex from $C$ whose $R$-neighbourhood $a$ lies has to be special). Similarly, every $b \in B$ whose negative $(r, q)$-type is not saturated is in $N_R(d'_i)$ and hence in $N_{2R}(d'_i)$ for some $i \in \{1, \ldots, m\}$.

Now it is easy to define a winning strategy for the duplicator in the $q$-round X-positive EF-game on $\mathcal{A}$ and $\mathcal{B}$: If the spoiler plays near a vertex $c'_i$ or $d'_i$, the duplicator answers according to the local strategy there. If the spoiler plays near a vertex played before, the duplicator answers according to the local strategy there. Otherwise, the spoiler plays a saturated vertex far away from everything, and the duplicator can always find an answer. The meaning of “near” varies with the number $j$ of moves remaining in the game. The duplicator seeks to preserve neighbourhoods of radius $3j$ around previously played elements or $2R + 3j$ around $c'_i$ or $d'_i$.

This finally completes the proof of Lemma 7 and thus, altogether, the proof of Theorem 2. □

4. Graph Decompositions

In this section we fix some notation and briefly present the basic notions from graph minor theory used later on. See the last chapter of [8] or the survey [24].

The vertex set of a graph $\mathcal{G}$ is denoted by $V^\mathcal{G}$ and its edge set is denoted by $E^\mathcal{G}$. For $U \subseteq V^{\mathcal{G}}$ we write $\langle U \rangle$ for the subgraph of $\mathcal{G}$ induced by $V$. A tree is an acyclic, connected graph. We usually use rooted directed trees where edges are directed from the root towards the leaves.

A minor of a graph $\mathcal{G}$ is a graph $\mathcal{H}$ that can be obtained from a subgraph of $\mathcal{G}$ by contracting edges. We write $\mathcal{H} \preceq \mathcal{G}$ to denote that $\mathcal{H}$ is a minor of $\mathcal{G}$. A class $\mathcal{C}$ of graphs is minor closed if, and only if, for all $\mathcal{G} \in \mathcal{C}$ and $\mathcal{H} \preceq \mathcal{G}$ also $\mathcal{H} \in \mathcal{C}$. A class $\mathcal{C}$ of graphs is $\mathcal{H}$-tree if $\mathcal{H} \not\preceq \mathcal{G}$ for all $\mathcal{G} \in \mathcal{C}$. We then call $\mathcal{H}$ an excluded minor of $\mathcal{C}$.

A tree-decomposition of a graph $\mathcal{G}$ is a pair $(T, (B_t)_{t \in V(T)})$, where $T$ is a directed tree and $B_t \subseteq V^\mathcal{G}$ for all $t \in V^T$ such that $\bigcup_{t \in V^T} B_t = \mathcal{G}$ and for every $v \in V^{\mathcal{G}}$ the set $\{t : v \in B_t\}$ is connected. The sets $B_t$ are called blocks of the decomposition. The width of $(T, (B_t)_{t \in V(T)})$ is $\max \{|B_t| : t \in V^T\} + 1$ and the tree-width $tw(\mathcal{G})$ of a graph $\mathcal{G}$ is the minimal width of any of its tree-decompositions. A class $\mathcal{C}$ of graphs has bounded tree-width, if there is a constant $k$ bounding the tree-width of all members of $\mathcal{G}$.

For a tree-decomposition $(T, (B_t)_{t \in V(T)})$ and $t \in V^T$ with parent $s \in V^T$ we let $A_t := B_t \cap B_s$. For the root $r$ of $T$ we let $A_r := \emptyset$. The adhesion of $(T, (B_t)_{t \in V(T)})$ is the number $\operatorname{ad}(T, (B_t)_{t \in V(T)}) := \max \{|A_t| : t \in V^T\}$. The torso $[B_t]_t$ of $(T, (B_t)_{t \in V(T)})$ at $t \in V^T$ is the graph with vertex set $B_t$ and with an edge between $u, v \in B_t$ if $(u, v) \in E^\mathcal{G}$ or $u, v$ both belong to a block $B_s$ with $s \neq t$.

A tree-decomposition of a graph $\mathcal{G}$ over a class $\mathcal{B}$ of graphs is a tree-decomposition $(T, (B_t)_{t \in V(T)})$ whose torsis $\{B_t\}$ are contained in $\mathcal{B}$.

We also need the following notion. The local tree-width $\operatorname{ltw}^\mathcal{G}$ of a graph $\mathcal{G}$ is the function $\operatorname{ltw}^\mathcal{G} : N \to N$ defined as $\operatorname{ltw}^\mathcal{G}(r) := \max \{|N_\mathcal{G}(v)| : v \in V^{\mathcal{G}}\}$. A class $\mathcal{C}$ of graphs has bounded local tree-width if there is a function $f : N \to N$ such that $\operatorname{ltw}^\mathcal{G}(r) \leq f(r)$ for all $\mathcal{G} \in \mathcal{C}$ and $r \in N$.

5. First-order definable optimisation problems

In this section we present a proof of Theorem 1. Here, we only prove the minimisation version of the theorem. The maximisation version is proved similarly using techniques from [14]. We defer the details to the full version of the paper. We begin with a formal definition. Let $X_{\text{min}}$ be an optimal solution for $\operatorname{Min}_{\varphi(X)}$ on input $\mathcal{G}$. For $\varepsilon > 0$ we call a solution $X$, i.e. a set $X$ with $\langle \mathcal{G}, X \rangle = \varphi$, $\varepsilon$-close if $|X| \leq (1 + \varepsilon)|X_{\text{min}}|$. A polynomial-time approximation scheme (PTAS) for $\operatorname{Min}_{\varphi(X)}$ is an uniform family $\{A_\varepsilon\}_{\varepsilon > 0}$ of algorithms, where $A_\varepsilon$ given an instance $\mathcal{G}$, computes an $\varepsilon$-close solution for $\mathcal{G}$ in polynomial time. Uniform here
means that there is an algorithm that, given \( \epsilon \), generates \( A_\epsilon \). A PTAS is called \textit{efficient}, (or, it is an EPTAS), if the degree of the polynomial bounding the running time of \( A_\epsilon \) does not depend on \( \epsilon \). Our proof of Theorem 1 establishes an EPTAS for first-order definable optimisation problems.

**Example 11.** It is well known that the class of planar graphs excludes a minor. Thus, by Theorem 1, every optimisation problem definable by an X-positive or X-negative first-order formula has a PTAS on the class of planar graphs.

However, the result neither extends to monadic second-order logic (MSO) nor to first-order formulae which are not monotone in \( X \). For this, note that 3-colourability is NP-complete even on the class of planar graphs (see [13]). As 3-colourability can easily be defined by a formula \( \psi \in \text{MSO} \), the minimisation problem defined by \( \varphi(X) := \psi \rightarrow \forall x X x \) cannot have a PTAS (unless \( P = \text{NP} \)). Similarly, a simple reduction shows that the 3-colourability problem on planar graphs can be reduced to a minimisation problem on planar graphs defined by a non-monotone first-order formula. \( \Box \)

To prove Theorem 1 we first need some preparation. Let \( \varphi \) be a first-order formula positive in \( X \). By Theorem 2 we can assume that \( \varphi := \bigvee_j \land_i \psi_{j,i} \), where each \( \psi_{j,i} \) is \( X \)-positive and either basic local or the negation of a basic local formula. To compute a minimal set \( \psi \rightarrow \varphi \) of a basic local formula. To compute a minimal set

\[
\psi \rightarrow \varphi
\]

of the polynomial bounding the running time of \( \psi \).

Proof of Theorem 1 establishes an EPTAS if the degree \( k \) of \( X \) is either an existential basic formula. Hence, Theorem 1 follows from the following lemma.

**Lemma 12.** For every set \( X \subseteq V^\varphi \), \( (\varphi,X) \models \psi \) if, and only if, \( (\varphi,X) \models \psi^* \).

Proof of Theorem 12 implies that for a given graph \( \varphi \) we can translate the formula \( \varphi(X) \) into a conjunction of formulae of the form \((a)\) and \((*)\). By distributivity, we can translate this into a disjunction of conjunctions of formulae \( \chi_{j} \pi_{f,n} \) and formulae \( \xi := \bigwedge_{j=1}^{k} \vartheta(b_i) \) for tuples of constants \( \pi, \pi' \), functions \( f \) and numbers \( j, n \). As the arity of the tuples \( \pi, \pi' \) is bounded by a function of \( \varphi \) and the ranges of \( j, n, f \) also only depend on \( \varphi \), the translation can be done in polynomial time in the size of \( \varphi \).

For the first line of the formula \( \chi_{j} \pi_{f,n} \pi \) only imposes conditions on the choice of the tuple \( \pi \). It follows that for computing an approximation of a set \( X \) satisfying \( \varphi \) in \( \varphi \) it suffices to compute an approximation of a set \( X \) satisfying the conjunction of formulae

\[
\bigwedge_{k=1}^{k} \vartheta(a_i)
\]

for an \( r \)-local formula \( \vartheta \) and a tuple of constants \((a_1,\ldots,a_k)\) with \( \text{dist}(a_i, a_j) > 2r \) for all \( i \neq j \).

\[
\bigwedge_{k=1}^{k} \vartheta(x_i)
\]

for an \( r \)-local formula \( \vartheta \) and a tuple of elements \((a_1,\ldots,a_k)\) of distance \( \text{dist}(a_i, a_j) > 2q \) for all \( i \neq j \).

\[
\bigwedge_{k=1}^{k} \vartheta(x_i)
\]

for an \( r \)-local formula \( \vartheta \) and a tuple of constants \((q_1,\ldots,q_k)\) with \( q_i > 2r \) only depending on \( \varphi \).

Note that the formulas in \((c)\) are \( 5^r \cdot 4r \)-local around \( a_i \).

Hence, Theorem 1 follows from the following lemma.

**Lemma 13.** Let \( \varphi \) be a class of graphs with an excluded minor and let \( \sigma := \{a_1,\ldots,a_k\} \) be a set of constant symbols. Let \( q > 0 \) and let \( \varphi(x) \in \text{FO} \) be an X-positive conjunction of \( q \)-local formulae \( \psi(a) \) using only one constant symbol \( a \) in \( \sigma \) and formulae \( \psi := \bigwedge_{k=1}^{k} \vartheta(x_i, a_{i,t}) \), for an \( r \)-local formula \( \vartheta \) with \( q_i > 2r \), using constant symbols \( \pi_1, \ldots, \pi_k \subseteq \sigma \). Let \( \varphi(X) := \bigwedge_{i=1}^{n} \vartheta(x_{i} \pi_{f,n} \pi_{f,n}) \) has a polynomial time approximation scheme.
To prove the lemma we use a decomposition theorem for classes of graphs with an excluded minor that is due to [14].

We first introduce some notation.

For $\lambda, \mu \geq 0$ we let

$$\mathcal{L}(\lambda) := \{\mathcal{G} : \text{for all } \mathcal{H} \leq \mathcal{G}, \text{for all } r \geq 0 \text{ ltw}_{\mathcal{H}}(r) \leq \lambda \cdot r\}$$

$$\mathcal{L}(\lambda, \mu) := \{\mathcal{G} : \text{there is } X \subseteq V(\mathcal{G}) \text{ s.th. } (|X| \leq \mu \land \mathcal{G}\backslash X \in \mathcal{L}(\lambda))\}$$

Note that $\mathcal{L}(\lambda, \mu)$ is minor closed. The proof of Lemma 13 is based on the following decomposition theorem for classes of graphs with an excluded minor.

**Theorem 14 ([14]).** Let $\mathcal{C}$ be a class of graphs with an excluded minor. Then there exist $\lambda, \mu \in \mathbb{N}$ such that all $\mathcal{G} \in \mathcal{C}$ have a tree-decomposition over $\mathcal{L}(\lambda, \mu)$.

For the approximation algorithms we want to show, we need an algorithmic version of this theorem.

**Theorem 15 ([14]).** Let $\mathcal{C}$ be a minor closed class of graphs. Then there is a polynomial-time algorithm that computes for a given graph $\mathcal{G}$ a tree-decomposition of $\mathcal{G}$ over $\mathcal{C}$ or rejects $\mathcal{G}$ if no such decomposition exists.

Let $\mathcal{G}$ be a graph. For every vertex $v \in V(\mathcal{G})$ and integers $j \geq 0$ we define

$$L^G_v[i, j] := \{w \in V(\mathcal{G}) : i \leq \text{dist}(v, w) \leq j\},$$

where $\text{dist}(v, w)$ denotes the distance between $v$ and $w$ in $\mathcal{G}$. To simplify notation, we will use $L^G[i, j]$ for arbitrary integers $i$ and $j$ and set $L^G[i, i] := \emptyset$ if $i > j$ and $L^G[i, i] := L^G[0, i]$ for $i \leq 0$. The following lemma follows easily.

**Lemma 16.** Let $\lambda \in \mathbb{N}$. Then $\text{tw}(L^G[i, j]) \leq \lambda \cdot (j - i + 1)$ for all $\mathcal{G} \in \mathcal{L}(\lambda)$, $v \in V(\mathcal{G})$ and $i, j \in \mathbb{Z}$ with $i \leq j$.

Now, let $\mathcal{G}$ be as in the hypothesis of Lemma 13 and let $\mathcal{C}$ be a class of graphs with an excluded minor. By Theorem 14, we can choose $\lambda, \mu \in \mathbb{N}$ such that every graph $\mathcal{G} \in \mathcal{C}$ has a tree-decomposition over $\mathcal{L}(\lambda, \mu)$. Let $\varepsilon > 0$. We describe a polynomial time algorithm that, on input $\mathcal{G} \in \mathcal{C}$ and $\pi \in V(\mathcal{G})$, computes an $\varepsilon$-close solution for $\text{MIN}_{\mathcal{G}(X)}(\mathcal{C})$. To ease notation we will consider the tuple $\pi$ as part of the graph and use notation such as $\mathcal{G} \models \phi$ for $\langle \mathcal{G}, \pi \rangle \models \phi$.

The proof of Lemma 13 is split into two steps. In the first step, which we present in the next subsection, we prove the lemma for the classes $\mathcal{L}(\lambda)$ and $\mathcal{L}(\lambda, \mu)$ of graphs. Here, we use the corresponding result for graphs of bounded tree-width which essentially follows from [2].

**Theorem 17 ([2]).** Let $\phi(X)$ be an $X$-positive formula of MSO. Then $\text{MIN}_{\phi(X)}(\mathcal{C})$ can be solved in linear time on any class $\mathcal{C}$ of graphs of bounded tree-width.

In Section 5.3, we extend the proof to graphs which have a tree-decomposition over $\mathcal{L}(\lambda, \mu)$, i.e. to all graphs in $\mathcal{C}$.

### 5.1. The levels of graphs of bounded local tree-width.

In the first step of the proof of Lemma 13 we show that the restriction of $\text{MIN}_{\phi(X)}(\mathcal{C})$ to instances in $\mathcal{L}(\lambda)$ has a PTAS. Let $\phi := \bigwedge_{i \in \mathbb{N}} \phi_i(\alpha_i) \land \bigwedge_{i \in \mathbb{N}} \psi_i$, where the $\phi_i(\alpha_i)$ are $r$-local formulae and the $\psi_i := \forall x(\bigwedge_{t=1}^{2q} \text{dist}(x, a_{i,t}) > r_i \rightarrow \sigma_i(x))$, for a $q$-local formula $\sigma_i$ with $r_i > 2q$, using constant symbols $\alpha_i \subseteq \sigma$. For simplicity we assume w.l.o.g. that $r_i = r_j$ for all $i, j \in \mathbb{N}$. Let $r := r_i$.

Let $k := \lceil \frac{2q}{r} \rceil$. Note that $\frac{2q}{r} \leq (1 + \varepsilon)$. For $1 \leq i < j$ and $j \geq 0$ let $L_{ij} := L^G[i - 1] \cdot j - r + 1$ and $L_{ij} := L^G[j + 2r + 1]$. By Lemma 16, $\text{tw}(LI_{ij}) \leq \lambda(k + 2r + 1).

For all $1 \leq i < j, j \geq 0$ let $X_{ij}$ be a set of minimal cardinality such that

1. $\langle LI_{ij}, X_{ij} \rangle \models \phi_i(\alpha_i)$ for all $l \in L_{ij}$ such that $\alpha_i$ and its $r$-neighbourhood is contained in $L_{ij}$ and
2. $\langle LI_{ij}, X_{ij} \rangle \models \psi_i$ for all $l \in L_{ij}$.

Note that as $\psi_i := \forall x(\bigwedge_{t=1}^{2q} \text{dist}(x, a_{i,t}) > r_i \rightarrow \sigma_i(x))$ also mentions constants interpreted by vertices outside of $L_{ij}$, this is, strictly speaking, not well defined. However, as the $\alpha_i$ are constants, we can easily check whether $x$ is close to any constant interpreted by an element outside of $L_{ij}$. For instance, we could colour the $r$-neighbours of $a_1, \ldots, a_{2q}$ and then check in $\psi_i$ that $x$ is outside a coloured area. For ease of presentation we will therefore simply write $\langle LI_{ij}, X_{ij} \rangle \models \psi_i$ even in cases where some or all of the constants are interpreted by elements outside of $L_{ij}$.

By Theorem 17 the sets $X_{ij}$ can be computed in linear time. For $1 \leq i \leq s$ let $X_i := \bigcup_{j \geq 0} X_{ij}$. As $\phi$ is monotone in $X$, every $X_i$ is a solution of $\text{MIN}_{\phi(X)}(\mathcal{C})$ on $\mathcal{G}$. Let $X_{\text{min}}$ be an optimal solution of $\text{MIN}_{\phi(X)}(\mathcal{C})$ for $\mathcal{G}$, i.e. a set of minimal cardinality such that $\langle \mathcal{G}, X_{\text{min}} \rangle \models \phi$. Clearly, $X_{\text{min}} \cap X_{ij}$ satisfies the conditions (1) and (2) above for all levels $L_{ij}$. Hence,

$$\sum_{i=1}^{k} |X_i| \leq k \sum_{i=1}^{k} |X_{ij}| \leq \sum_{i=1}^{k} |L_{ij} \cap X_{\text{min}}| \leq (k + 2r)|X_{\text{min}}|.$$
Hence, $X_m$ is an $\varepsilon$-close solution of $\text{MIN}_{\phi(X)}(\mathcal{G})$ on $\mathcal{G}$. As every set $X$ can be computed in polynomial time, the set $X_m$ can also be computed in polynomial time.

5.2. Extension to the classes $\mathcal{L}(\lambda, \mu)$.

In a second step we show how this approximation algorithm can be extended to the classes $\mathcal{L}(\lambda, \mu)$ for constants $\lambda, \mu \geq 0$. Let $\mathcal{G} \in \mathcal{L}(\lambda, \mu)$ and let $U \subseteq V^d$ be such that $|U| \leq \mu$ and $\mathcal{G} \setminus U \in L(\lambda)$. The following extension of Theorem 17 can easily be proved by dynamic programming.

**Theorem 18.** For every $k \geq 0$ and every first-order formula $\phi(X)$ which is positive in the set-variable $X$, the following problem can be solved in linear time. Given a graph $\mathcal{G}$, a set $U \subseteq V^d$ so that $tw(\mathcal{G} \setminus U) \leq k$ and a subset $Y \subseteq U$, find a set $X \subseteq V^d \setminus U$ of minimal cardinality such that $(\mathcal{G}, X \cup Y) \models \phi$ or determine that no such set exists.

Let again $\phi(X)$ be an $X$-positive first-order formula. For every $Y \subseteq U$ let $X(Y)$ be a subset of $V^d \setminus U$ such that $(\mathcal{G}, X(Y) \cup Y) \models \phi$ and $|X(Y)| \leq (1 + \varepsilon) \min\{|X : X \subseteq V^d \setminus U \text{ and } (\mathcal{G}, X \cup Y) \models \phi\}$ or $X(Y) := \bot$ if no such set exists. If $X(Y) := \bot$ for all $Y \subseteq U$ then $\text{MIN}_{\phi(X)}(\mathcal{G})$ has no solution on $\mathcal{G}$ and we are done. Otherwise let $Y_0 \subseteq U$ be such that $|X(Y_0) \cup Y_0|$ is minimal among $|X(Y) \cup Y| : Y \subseteq U$ and $X(Y) \neq \bot$. Then clearly, $X(Y_0) \cup Y_0$ is an $\varepsilon$-close solution for $\text{MIN}_{\phi(X)}(\mathcal{G})$.

Using Theorem 18 instead of Theorem 17, the sets $X(Y)$ can be computed in polynomial time analogously to the first step. As there are only $2^u$ possible subsets of $U$ – recall that $\mu$ is a constant only depending on the class $\mathcal{G}$ – and for each $Y \subseteq U$, $X(Y)$ can be computed in polynomial time, the solution $X(Y_0) \cup Y_0$ can be computed in polynomial time.

5.3. Excluded Minors.

In the last step, we show how the approximation algorithm can be extended to graphs that have a tree-decomposition over $\mathcal{L}(\lambda, \mu)$, i.e. to all graphs in $\mathcal{G}$.

Let $\mathcal{G} \in \mathcal{G}$. We first compute a tree-decomposition $(T, (B_t)_{t \in V^T})$ over $\mathcal{L}(\lambda, \mu)$. By Theorem 15, this can be done in polynomial time. Let $r$ be the root of $T$ and for every $t \in V^T$ with parent $s$ let $A_s := B_s \cap B_t$. We set $A_r := \emptyset$. Further, for every node $t \in V^T$ let $T_t$ be the subtree of $T$ rooted at $t$ and let $B_t := \bigcup_{i \in T_t} B_s$.

In what follows we will construct for subgraphs $\mathcal{B}$ of $\mathcal{G}$ sets $X$ such that $(\mathcal{B}, X)$ satisfies

1. $(\mathcal{B}, X) \models \phi(a_t)$ for all $t \in I$, such that $a_t$ and its $r$-neighbourhood is contained in $\mathcal{B}$ and

2. $(\mathcal{B}, X) \models \psi_t$ for all $t \in I$.

(Here, we use the same convention as in Section 5.1 above.) To simplify the presentation we write $(\mathcal{B}, X) \models \phi$ to indicate that $X$ satisfies the conditions (1) and (2) in $\mathcal{B}$. The notation is motivated by the fact that for $\mathcal{B} = \mathcal{G}$, $(\mathcal{G}, X) \models \phi$ for any set satisfying condition (1) and (2) and vice versa.

Inductively, from the leaves to the root, we compute for every node $t \in V^T$ and for every subset $Y \subseteq A_t$ an $X(t, Y)$ such that $X(t, Y) \subseteq \mathcal{B} \setminus A_t$, $(\mathcal{B}, X(t, Y) \cup Y) \models \phi$ and $|X(t, Y)| \leq (1 + \varepsilon) \min\{|X : (\mathcal{B}, X \cup Y) \models \phi, X \subseteq B_t \setminus A_t\}$ or $X(t, Y) := \bot$ if no such set exists. As tree-decompositions over $\mathcal{L}(\lambda, \mu)$ have adhesion at most $\lambda + \mu + 1$, we have $|A_t| \leq \lambda + \mu + 1$. Hence, we only have to compute a constant number of sets $X(t, Y)$ for each $t$. Further, for the root $r$ we have $A_r = \emptyset$ and $(\mathcal{B}_r) = \mathcal{G}$. Hence, $(\mathcal{G}, \emptyset)$ is an $\varepsilon$-close solution for $\text{MIN}_{\phi(X)}(\mathcal{G})$ or if no solution exists.

We show next how to compute the sets $X(t, Y)$. Suppose $t \in V^T$ and for every child $t'$ of $t$ we have already computed the family $X(t', \cdot)$. Let $U \subseteq B_t$ such that $|U| \leq \mu$ and $|B_t \setminus U| \leq \mathcal{L}(\lambda)$. (Recall that $|B_t|$ denotes the torsor of $(T, (B_t)_{t \in V^T})$ at $t$.) Let $W := U \cup A_t$.

For every $Z \subseteq W$, let $X_{\text{min}}(Z)$ be a set of minimal cardinality such that $X_{\text{min}}(Z) \subseteq B_t \setminus W$ and $(\mathcal{B}_t, X_{\text{min}}(Z) \cup Z) \models \phi$ or $X_{\text{min}}(Z) := \bot$ if no such set exists.

**Claim 19.** For every set $Z \subseteq W$ we can compute in polynomial time an $X(t, Y)$ such that $X(t, Y) \subseteq B_t \setminus W$, $(\mathcal{B}_t, X(t, Y) \cup Z) \models \phi$, and $|X(t, Y)| \leq (1 + \varepsilon)|X_{\text{min}}(Z)|$ or $X(t, Y) := \bot$ if no such set exists.

Before we prove Claim 19 let us show how the proof of Lemma 13 can be completed using the claim. For every $Y \subseteq A_t$ choose a $Z \subseteq W$ such that $Z \cap A_t = Y$ and $|X(Z) \cup Z| = \min\{|X(Y) \cup Z : Y \subseteq Z \subseteq W, Z \cap A_t = Y\}$. Set $X(t, Y) := X(Z) \cup Z \cap A_t$. By our choice of $Z$ it follows that $|X(t, Y)| \leq (1 + \varepsilon)|X_{\text{min}}(Z)|$ which concludes the proof.

So all that remains is to prove Claim 19. Fix a $Z \subseteq W$. We show how to compute $X(t, Y)$ in polynomial time. If $W = B_t$, i.e. $B_t := U \cup A_t$, then let $X(t, Y) := X_{\text{min}}(t, Y) \cup A_t$. Otherwise choose an arbitrary vertex $v \in B_t \setminus W$. For $1 \leq i \leq k$ and $j \geq 0$ let $L_{ij} := L_{ij}^0 \setminus W \setminus \{(j - 1) \cdot k + i - r, j \cdot k + i + r\}$. Then tw($L_{ij}$) $\leq \lambda (k + 1 + 2r)$. For every child $t'$ of $t$ and every $1 \leq i \leq k$ there is at least one $j \geq 0$ such that $A_t \setminus \bigcup_{i \leq k} L_{ij}$. This follows from that fact that $A_t$ induces a clique in $[B_t]$. Let $j_{\text{min}}(t, t')$ be the least such $j$ and let $L_{i,j} := L_{ij} \cup \bigcup_{t \in L_{ij}} (\mathcal{B}_t \setminus A_t) : (t, t') \in E^T$ and $j_{\text{min}}(t, t') = j$.

Similarly, for every $X \subseteq L_{ij}$ let $X^* := X \cup \bigcup_{t \in L_{ij} \setminus X} (X(t', X \cup Z) \cup A_t) : (t, t') \in E^T, j_{\text{min}}(t, t') = j$.

We compute an $X_{\text{ij}} \subseteq L_{ij}$ with minimal $|X_{\text{ij}}|$ such that $(L_{ij}, X_{\text{ij}} \cup Z) \models \phi$ or set $X_{\text{ij}} := \bot$ if no such $X$ exists. This can be done in polynomial time using the standard dynamic programming techniques on graphs of bounded tree-width.
provided that the numbers \(|X(t', Y)|\) for the children \(t'\) of \(t\) are given. It is important here that every \(A_r \setminus W\) is a clique in \((L_{ij})\), as this implies that it is contained in a single block of every tree-decomposition of \((L_{ij})\).

Let \(X_i := \bigcup_{j=0}^l X_{ij}\) and \(X_i^* := \bigcup_{j=0}^l X_{ij}^*\). Then, by monotonicity of \(\varphi\) in \(X, ((R_i), X_i^* \cup Z) \models \varphi\) or \(X_i = \perp\) if no set satisfying \(\varphi\) in \((R_i)\) exists. Finally, choose an \(i \in \{1, \ldots, k\}\) such that \(|X_i^*| = \min\{|X_1^*|, \ldots, |X_k^*|\}\) and let \(X(Z) := X_i^*\). It follows that \(|X(Z)|\) can be computed in polynomial time.

Recall that we defined \(X_{\min} := X_{\min}(Z) \subseteq R_i \setminus W\) to be a set of minimal order such that \(((R_i), X_{\min} \cup Z) \models \varphi\) or \(X_{\min} := \perp\) if no such set exists. It remains to prove that \(|X(Z)| \leq (1 + \varepsilon)|X_{\min}|\).

By hypothesis of the algorithm we have for every child \(t'\) of \(t\), \(|X(t', (X_{\min} \cup Z) \cap A_r)| \leq (1 + \varepsilon)|X_{\min} \cap (R_i \setminus A_r)|\).

Further, the construction of \(X_j\) and \(X_{ij}\) guarantees that for \(1 \leq i \leq k\) and \(j \geq 0\):

\[|X_{ij}| \leq |X_{\min} \cap L_{ij}| + \sum_{(i,j) < (i',j')} |X(t', (X_{\min} \cup Z) \cap A_r)|,\]

But then

\[k|X(Z)| \leq \sum_{i=1}^k |X_i^*| \leq \sum_{i=1}^k \sum_{j=0}^l |X_{ij}|\]
\[\leq \sum_{i=1}^k \sum_{j=0}^l \left(|X_{\min} \cap L_{ij}| + \sum_{(i',j') < (i,j)} |X(t', (X_{\min} \cup Z) \cap A_r)|\right)\]
\[\leq \sum_{i=1}^k \sum_{j=0}^l \left(|X_{\min} \cap L_{ij}| + \sum_{(i',j') < (i,j)} (1 + \varepsilon)|X_{\min} \cap (R_i \setminus A_r)|\right)\]
\[\leq (k + 2r)|X_{\min} \cap B_i| + k(1 + \varepsilon)|X_{\min} \cap (R_i \setminus B_i)|\]

This implies \(|X(Z)| \leq (1 + \varepsilon)|X_{\min}|\) and concludes the proof of Lemma 13 and with it also the proof of Theorem 1.

References

APPENDIX

A. Proofs omitted in Section 3

Proof of Theorem 2:
Let $\xi$ be an $X$-positive FOF$(\tau \cup \{X\})$-sentence, let $q$ be the quantifier rank of $\xi$, and let $K, R, Q \geq 2$ be chosen according to Lemma 5. It should be clear that, up to logical equivalence, each of the sets $BL^+(K, R, Q)$ and $BL^-(K, R, Q)$ only contains a finite number of distinct sentences. Let $M > 0$ be an upper bound on this number and let $\varphi_1^+, \ldots, \varphi_M^+$ (respectively, $\varphi_1^-, \ldots, \varphi_M^-$) be a list of all (up to logical equivalence) sentences in $BL^+(K, R, Q)$ (respectively, $BL^-(K, R, Q)$).

For sets $I^+, I^- \subseteq \{1, \ldots, M\}$ we let

$$\theta_{I^+, I^-} := \bigwedge_{i \in I^+} \varphi_i^+ \land \bigwedge_{j \in I^-} \neg \varphi_j^-.$$ 

For every $\tau \cup \{X\}$-structure $\alpha$ we let

$$I^+_\alpha := \{i \in \{1, \ldots, M\} : \alpha \models \varphi_i^+\} \quad \text{and} \quad I^-_\alpha := \{j \in \{1, \ldots, M\} : \alpha \models \neg \varphi_j^-\}.$$ 

Obviously, $\theta_{I^+, I^-}_\alpha$ is an $X$-positive sentence such that $\alpha \models \theta_{I^+, I^-}_\alpha$.

We let $\xi'$ be the disjunction of the sentences $\theta_{I^+, I^-}$, for all those $I^+, I^- \subseteq \{1, \ldots, M\}$, for which there exists a $\tau \cup \{X\}$-structure $\alpha$ with $\alpha \models \xi$, $I^+_\alpha = I^+$, and $I^-_\alpha = I^-$. Obviously, $\xi'$ is an $X$-positive Boolean combination of (symmetric) basic local sentences. To complete the proof of Theorem 2, it remains to show that $\xi'$ is equivalent to $\xi$.

To this end, let $\mathcal{B}$ be an arbitrary $\tau \cup \{X\}$-structure. Our aim is to show that $\mathcal{B} \models \xi' \iff \mathcal{B} \models \xi$.

For the direction $\Rightarrow$, note that if $\mathcal{B} \models \xi'$, then $\theta_{I^+, I^-}_\mathcal{B}$ is one of the disjuncts in $\xi'$, and thus $\mathcal{B} \models \xi'$.

For the direction $\Leftarrow$, note that if $\mathcal{B} \models \xi'$, then $\mathcal{B}$ satisfies (at least) one of the disjuncts of $\xi'$. I.e., there is a $\tau \cup \{X\}$-structure $\alpha$ with $\alpha \models \xi'$ such that $\mathcal{B} \models \theta_{I^+, I^-}_\alpha$.

Since $\mathcal{B} \models \theta_{I^+, I^-}_\alpha$, we know that every $\varphi \in BL^+(K, R, Q)$ holds in $\alpha$ also holds in $\mathcal{B}$ and every $\varphi \in BL^-(K, R, Q)$ that holds in $\mathcal{B}$ also holds in $\alpha$. Therefore, by applying Lemma 5, we obtain that $\alpha \models \xi$ implies $\mathcal{B} \models \xi$. This completes the proof of Theorem 2. \qed

Proof of Proposition 6:
We prove the contrapositive. I.e., assuming that there is an $X$-positive FOF$(\tau \cup \{X\})$-sentence $\xi$ of quantifier rank at most $q-i$ such that $(\mathcal{A}, a_1, \ldots, a_i) \models \xi$ and $(\mathcal{B}, b_1, \ldots, b_i) \models \neg \xi$. Note that, in particular, for $i = q$ this implies that the mapping $(a_1 \mapsto b_1)_{1 \ldots q}$ cannot be an $X$-positive partial isomorphism from $\mathcal{A}$ to $\mathcal{B}$, and thus the spoiler will have won the game.

The basis $i = 0$ is trivially satisfied with $\xi' := \xi$. For the induction step we know that there is an $X$-positive formula $\xi'_i$ of quantifier rank at most $q-i$ such that $(\mathcal{A}, a_1, \ldots, a_i) \models \xi_i$ and $(\mathcal{B}, b_1, \ldots, b_i) \models \neg \xi'_i$. If $\xi'_i$ is of the form $\exists x \chi$ (respectively, $\forall x \chi$), then letting $\xi'_{i+1} := \chi$, the spoiler can choose an element $a_{i+1}$ in $\mathcal{A}$ (respectively, $b_{i+1}$ in $\mathcal{B}$) such that $(\mathcal{A}, a_1, \ldots, a_{i+1}) \models \xi'_{i+1}$ (respectively, $(\mathcal{B}, b_1, \ldots, b_{i+1}) \models \neg \xi'_{i+1}$), whereas for any answer $b_{i+1}$ (respectively, $a_{i+1}$) that the duplicator may give, we have $(\mathcal{B}, b_1, \ldots, b_{i+1}) \models \neg \xi'_{i+1}$ (respectively, $(\mathcal{A}, a_1, \ldots, a_{i+1}) \models \xi'_{i+1}$). In the remaining cases, $\xi'_i$ does not start with a quantifier, but it is straightforward to find a subformula $\xi''_i$ of $\xi'_i$ that (1) starts with a (universal or existential) quantifier, (2) has the same free variables as $\xi_i$, and (3) satisfies $(\mathcal{A}, a_1, \ldots, a_i) \models \xi''_i$ and $(\mathcal{B}, b_1, \ldots, b_i) \models \neg \xi''_i$. Now, letting $\xi'_{i+1} := \exists x \chi$ or $\forall x \chi$, one can proceed as in the previous case. \qed

Proof of Lemma 8:
Let $k, r, q \geq 0$ and let $\mathcal{A}, \mathcal{B}$ be $\tau \cup \{X\}$-structures. For some $m \geq 0$ only depending on $r, q, \mathcal{A}, \mathcal{B}$, and for $0 \leq i \leq m$, we inductively construct numbers $R_i, K_i, \kappa_i := 2k_i^2 - 1$, sets $C_i^\mathcal{A} \subseteq \mathcal{A}, C_i^\mathcal{B} \subseteq \mathcal{B}$, and sets $\Gamma_i^\mathcal{A}, \Gamma_i^\mathcal{B} \subseteq \Theta(r, q)$ of $(r,q)$-types such that:

(i) $R_m, K_m, \kappa_m, C_m^\mathcal{A}, C_m^\mathcal{B}$ satisfy (1)–(3);
(ii) $R_i, K_i, \kappa_i, C_i^\mathcal{A}, C_i^\mathcal{B}$ satisfy (1) and (2), for all $0 \leq i \leq m$;
(iii) $N_{R_{i+1}}(C_{i+1}^\mathcal{A}) \subseteq N_{R_i}(C_i^\mathcal{A})$, for each $\mathcal{D} \in \{\mathcal{A}, \mathcal{B}\}$ and for all $1 \leq i \leq m$;
(iv) each $\Theta \in \Gamma_i^\mathcal{B}$ is $R_i$-covered by $C_i^\mathcal{B}$, for each $\mathcal{D} \in \{\mathcal{A}, \mathcal{B}\}$ and for all $0 \leq i \leq m$;
(v) for all $1 \leq i \leq m$, the following is true: for each $\mathcal{D} \in \{\mathcal{A}, \mathcal{B}\}$, $\Gamma_i^\mathcal{D} \subseteq \Gamma_{i+1}^\mathcal{D}$, and for at least one $\mathcal{D} \in \{\mathcal{A}, \mathcal{B}\}$, the inclusion is strict.

By (v), $m$ is bounded by $2|\Theta(r,q)|$. Thus the construction terminates in a number of steps that only depends on $r$ and $q$ (which implies that we obtain upper bounds $\tilde{K}, \tilde{R}$ on $K_m, R_m$ that only depend on $r, q$, but not on the particular structures $\mathcal{A}, \mathcal{B}$).

As the induction basis, we let $R_0 = r$, $K_0 = k$, $\kappa_0 := 2k_0^2 - 1$, $C_0^\mathcal{A} = C_0^\mathcal{B} = \emptyset$, and $\Gamma_0^\mathcal{A} = \Gamma_0^\mathcal{B} = \emptyset$. Obviously, (ii)–(v) are satisfied.

For the induction step from $i$ to $i+1$, suppose that $R_i, K_i, \kappa_i := 2k_i^2 - 1, C_i^\mathcal{A}, C_i^\mathcal{B}, \Gamma_i^\mathcal{A}, \Gamma_i^\mathcal{B}$ satisfy (ii)–(iv). If $R_i, K_i, \kappa_i, C_i^\mathcal{A}, C_i^\mathcal{B}$ satisfy (1)–(3), then we let $m = i$ and we
are done. Assume otherwise. Then, there is $D \in \{\mathcal{A}, \mathcal{B}\}$ and $\theta \in \Theta(r, \bar{q})$ such that $\theta$ is neither $R_i$-covered by $C_i$ nor $(\kappa; 10R_i, 10K_i)$-free over $C_i$. Let us assume w.l.o.g. that $D = \mathcal{A}$ (the case where $D = \mathcal{B}$ is symmetric). Then, we inductively define elements $a_1, \ldots, a_t \in A$ as follows: If $\theta$ is $\kappa; 10R_i$-covered by $C_i$, then we let $t := 0$ (i.e., $\{a_1, \ldots, a_t\} = \emptyset$). Otherwise, there must be an element $a_1 \notin N_{\kappa; 10R_i}(C_i)$ such that $a_1$ realises type $\theta$. Now suppose that $a_1, \ldots, a_t$ are already defined. If $\theta$ is $\kappa; 10R_i$-covered by $C_i \cup \{a_1, \ldots, a_t\}$, then we let $t := j$ and stop. Otherwise, we let $a_{j+1} \notin N_{\kappa; 10R_i}(C_i \cup \{a_1, \ldots, a_t\})$ such that $a_{j+1}$ realises type $\theta$. The construction stops in at most $(10K_i - 1)$ steps, because otherwise $a_1, \ldots, a_{10K_i}$ would witness that $\theta$ is $(\kappa; 10R_i, 10K_i)$-free over $C_i$. Thus $t \leq 10K_i - 1$. Let $\tilde{C}_i = C_i \cup \{a_1, \ldots, a_t\}$ and note that $\theta$ is $\kappa; 10R_i$-covered by $\tilde{C}_i$. We let $\tilde{C}_i := C_i$. Furthermore, we let

$$K_{i+1} := 11K_i - 1$$

(thus, $|\tilde{C}_i|, |\tilde{C}_i| \leq K_{i+1}$),

$$K_2 = 2K_1 - 1,$$

$$\Gamma_i := \Gamma_i \cup \{\theta\},$$

$$\Gamma_1 := \Gamma_1.$$

We would like to let $R_{i+1} = \kappa; 10R_i$ and $C_{i+1} = \tilde{C}_i$, $\tilde{C}_i = \tilde{C}_i$, but that does not work because (2) is not necessarily satisfied. Instead, we inductively construct a sequence of radii $p_0 \leq \ldots \leq p_t$ and sets $D_0 \supseteq \ldots \supseteq D_t$, and then let $R_{i+1} = p_t$, $C_{i+1} = D_t$, and $\Gamma_{i+1} := D_t$. We start with $p_0 := \kappa; 10R_i$, $D_0 := \tilde{C}_i$, and $D_t := \Gamma_t$. Now suppose that $p_j$ and $D_j$ are already defined. If dist$(c, c') > \kappa; 10R_j$ for all $c, c'$ with $c \neq c'$ and $c, c' \in D_j$, we let $s = j$ and stop. Otherwise, there is $D' \in \{\mathcal{A}, \mathcal{B}\}$ and $c, c' \in D_j$ with $c \neq c'$ yet dist$(c, c') \leq \kappa; 10R_j$. We let $p_{j+1} := \kappa; 11R_j$, $D_{j+1} = D_j \setminus \{c\}$, and if $D' = \mathcal{A}$ then $D_{j+1} := D_j$, respectively, if $D' = \mathcal{B}$ then $D_{j+1} := D_j$.

This completes the construction. Note that in each step one element is removed from $D_j \cup D_j$, and thus the construction ends after at most $s \leq |\tilde{C}_i| + |\tilde{C}_i| \leq 2K_{i+1}$ steps.

We let

$$R_{i+1} := p_t,$$  

$$C_{i+1} := D_t,$$  

$$\Gamma_{i+1} := D_t.$$

We now prove that these numbers and sets satisfy (ii)–(v). For (ii), note that for each $D' \in \{\mathcal{A}, \mathcal{B}\}$,

- $|C_{i+1}| \leq K_{i+1}$, since $C_{i+1} = D_t \subseteq \tilde{C}_i$, and $|\tilde{C}_i| \leq K_{i+1}$;

- dist$(c, c') > \kappa; 10R_{i+1}$ for all $c, c' \in C_{i+1}$ with $c \neq c'$, by the construction of $D_t$.

For (iii) and (iv), note that for $0 \leq j < s$ and each $D' \in \{\mathcal{A}, \mathcal{B}\}$ we have $N_{p_j}(D_j) \subseteq N_{p_{j+1}}(D_j)$.

$$N_{R_i}(C_i) \subseteq N_{10R_i}(C_i) \subseteq N_{p_i}(D_i) \subseteq N_{p_{i+1}}(D_i) \subseteq N_{R_{i+1}}(C_{i+1}).$$

This directly proves (iii), and it implies (iv) as follows: Each type $\theta' \in \Gamma_i$ is $R_i$-covered by $\tilde{C}_i$ and hence $R_{i+1}$-covered by $\tilde{C}_i$. The type $\theta$ is $\kappa; 10R_i$-covered by $\tilde{C}_i$ and hence $R_{i+1}$-covered by $\tilde{C}_i$. Finally, (v) is trivially satisfied. Thus, the proof of Lemma 8 is complete.

**Proof of Lemma 10:**

Let $(N(X))$ denote the set of neighbours of a set $X$ of vertices of $G$. By Hall’s theorem applied to the induced subgraph of $G$ with vertex set $G \cup (N(G))$, there is a matching $M$ that covers each vertex of $N(G)$. A matching $M$ that covers all vertices in $G$ such that the number $k$ of vertices of $D_k$ that are not covered by $M$ is minimum. If $k = 0$, then all vertices of $G$ are covered by $M$, and the lemma is proved. Assuming for contradiction that $k > 0$. Let $d_0 \in D_k$ be a vertex that is not covered by $M$. By Hall’s theorem applied to the induced subgraph of $G$ with vertex set $D_k \cup (N(D_k))$, there is a matching $N$ that covers each vertex of $D_k$. Let us choose such a matching $N$. There is a unique maximal path $P$ starting in $d_0$ with edges alternating between $N$ and $M$. This path $P$ must end in a vertex not in $G \cup (N(D_k))$. Let $L := (G \setminus P) \cup (N \setminus P)$ be the edge set obtained from $M$ by removing all paths of $P$ and adding all $N$-edges. Then $L$ is also a matching that covers all vertices of $G$. Moreover, $L$ covers all vertices of $D_k$ covered by $M$ and in addition $d_0$. This is a contradiction to the minimality of $M$ and thus concludes the proof of Lemma 10.

**B. Proofs omitted in Section 5**

**Proof of Lemma 12:**

Suppose first that $(G, X) \models \psi$. Let $B := (b_1, \ldots, b_l) \subseteq V$ be a set of maximal cardinality such that $(G, X) \models \neg \theta(b_i)$ and dist$(b_i, b_j) > 2r$ for all $i \neq j$. Note that, by maximality of $B$, every $c \in V \cap \theta(c)$ belongs to $N_{2r}(B)$. As $(G, X) \models \psi$ we have $l < k$.

We construct a set $A \subseteq B$ and a radius $p \geq 8r$ such that

- (a) for all $b \in B$ there is an $a \in A$ with dist$(a, b) < \frac{p}{2}$
- (b) dist$(a, a') > p$ for all $a \neq a' \in A$.

For this, we inductively define a sequence of sets $A_n \subseteq B$ and radii $p_n$ such that $p_n < p_{n+1}$ and $A_n \supseteq A_{n+1}$ for all $n$ and every set $A_n$ satisfies condition (a) above.
Set \( p_0 := 8r \) and \( A_0 := B \). For the induction step suppose \( p_n \) and \( A_n \) are already defined. If \( \text{dist}(a, a') > p_n \) for all \( a \neq a' \in A \) we set \( p := p_n, A := A_n \) and terminate the induction. Otherwise there must be two elements \( a, a' \in A \) with \( a \neq a' \) and \( \text{dist}(a, a') \leq p_n \). Let \( p_{n+1} := 5p_n \) and \( A_{n+1} := A_n \setminus \{a'\} \).

By induction hypothesis, for every \( b \in B \) there is an \( a_b \in A_n \) with \( \text{dist}(a_b, b) < \frac{5p_n}{4} \). If \( a_b \neq a' \) then condition (a) remains satisfied for \( b \) in \( A_{n+1} \). Otherwise, \( \text{dist}(a', b) < \frac{5p_n}{4} \). Hence, \( \text{dist}(a, b) < \text{dist}(a, a') + \text{dist}(a', b) < p_n + \frac{5p_n}{4} \leq \frac{9p_n}{4} = \frac{p_{n+1}}{4} \).

So \( A_{n+1} \) satisfies condition (a).

Note that for all \( i, p_i = 5^i \cdot 8r \). Clearly, this process must stop after at most \( n \leq l < k \) induction steps resulting in a radius \( p := p_n \) and a set \( A := A_n \subseteq B \) satisfying both conditions (a) and (b).

Define \( j := |A| \) and let \( \mathcal{P} := \{a_1, \ldots, a_j\} \) list the elements of \( A := A_n \). Further, let \( P_l := \{b \in B : \text{dist}(b, a_i) < \frac{p}{4}\} \). Condition (a) implies that \( \bigcup P_l = B \). Let \( f \in \mathcal{P} \) be defined as \( f(i) := |P_i| \). We claim that \( (\mathcal{G}, X) \models \chi_{\mathcal{P}, f,n} \).

Towards a contradiction, suppose \( (\mathcal{G}, X) \not\models \chi_{\mathcal{P}, f,n} \). By construction, \( \text{dist}(a, a') > p = p_n = 5^i \cdot 8r \) for all \( a, a' \in A \). Hence, the conjunction in the first line of the definition of \( \chi_{\mathcal{P}, f,n} \) is satisfied. Further, \( (\mathcal{G}, X) \models \vartheta(c) \) for every element \( c \in V^\mathcal{G} \) such that \( \text{dist}(c, a) \geq 5^i \cdot 4r = \frac{5p_n}{4} \). For, as every element \( b \in B \) has distance less than \( \frac{5p_n}{4} \) from some \( a \in A \), \( \text{dist}(c, b) > \frac{5p_n}{4} - \frac{5p_n}{4} = 5^i \cdot 4r - 5^i \cdot 2r \geq 2r \). Hence, if there was such an element \( c \) with \( (\mathcal{G}, X) \models \neg \vartheta(c) \) this would contradict the maximality assumption for \( B \). So the conjunct in the second line of the definition of \( \chi_{\mathcal{P}, f,n} \) is satisfied as well. Thus, as \( (\mathcal{G}, X) \not\models \chi_{\mathcal{P}, f,n} \) there must be an element \( a_s \in A \) and elements \( c_1, \ldots, c_{f(s)+1} \) such that \( \text{dist}(c_i, c_{i+1}) > 2r \) for all \( i \neq s \) and for all \( i, (\mathcal{G}, X) \not\models \vartheta(c_i) \) and \( \text{dist}(c_i, a_s) < \frac{5p_n}{4} \). As \( \text{dist}(b, a_i) < \frac{5p_n}{4} \) for all \( b \in P_l, l \neq s \) it follows that for all \( c_i \) and all \( b \in P_l, l \neq s, \text{dist}(c_i, b) > 2r \).

But then \( B' := \{c_1, \ldots, c_{f(s)+1}\} \cup \bigcup_{l \neq s} P_l \) would contradict the maximality assumption of \( B \) as, by definition of \( f, f(s) + 1 > |P_s| \) and therefore \( |B'| > |B| \) but \( (\mathcal{G}, X) \not\models \vartheta(b') \) for all \( b' \in B' \) and also \( \text{dist}(b, b') > 2r \) for all \( b \neq b' \in B' \).

This concludes the first direction of the proof.

Now suppose \( (\mathcal{G}, X) \models \psi^* \), e.g. \( (\mathcal{G}, X) \models \chi_{\mathcal{P}, f,n} \) for some \( j, \mathcal{P}, f, n \), but \( (\mathcal{G}, X) \not\models \psi \). Then there must be a set \( B := \{b_1, \ldots, b_k\} \subseteq V^\mathcal{G} \) such that \( (\mathcal{G}, X) \models \neg \vartheta(b_i) \) and \( \text{dist}(b_i, b_j) > 2r \) for all \( i \neq j \). By definition of \( \chi_{\mathcal{P}, f,n} \) all \( b \in B \) must be within the \( 5^n \cdot 4r \)-neighbourhood of some \( a_i \). But, again by definition of \( \chi_{\mathcal{P}, f,n} \), in the \( 5^n \cdot 4r \)-neighbourhood around any \( a_i \), there can only be at most \( f(i) \) elements which are pairwise more than \( 2r \) apart and all satisfy \( \neg \vartheta \). As \( \sum_i f(i) < k \) we get a contradiction. \( \square \)